

GenCore version 5.1.4 p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 30, 2003, 01:24:25 ; Search time 10977.4 Seconds
(without alignments)
2152.747 Million cell updates/sec

Title: US-09-873-409-2
Perfect score: 4079
Sequence: 1 MVDENDIRALNVRHYRDHIG.....QELLNRDIYFKLVNAQSVQ 812

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2.1/USPTO.spool/US09873409/runat_27032003.115418.19228/app query fasta_1.7544
-DB=GenEmbl -QWTF=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09873409@cgn 1 1 54887 @runat_27032003.115418.19228 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*
1: gb.ba.*
2: gb.hcg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4079	100.0	2856	6 AX339028	AX339028 Sequence
2	4079	100.0	3177	6 AX339030	AX339030 Sequence
3	4079	100.0	3621	6 AX339032	AX339032 Sequence
4	4079	100.0	3702	6 AX339031	AX339031 Sequence
5	3919	96.1	3699	6 AX478104	AX478104 Sequence
6	3456	84.7	2066	6 AX339027	AX339027 Sequence
7	2480	60.8	4018	5 GGA9799	AJ009799 Gallus ga
8	2476	60.7	3987	10 CRUGP1185	M59254 Chinese ham
9	2476	60.7	4296	10 CRUGP1	M60040 C.griseus p
10	2476	60.7	4304	10 CRUGP1165	M59253 Chinese ham
11	2468	60.5	3858	4 OAU78609	U78609 Ovis aries
12	2467	60.5	4084	10 MUSMDRA	J03398 Mouse mdr g
13	2465	60.4	3924	6 AX024455	X024455 Sequence
14	2465	60.4	3924	6 AX409652	AX409652 Sequence
15	2465	60.4	3924	9 HUMMDR3	M23234 Human membr
16	2451	60.1	4927	10 AF257746	AF257746 Rattus no
17	2442	59.9	4279	6 AX105057	AX105057 Sequence
18	2442	59.9	4279	6 AX105078	AX105078 Sequence
19	2442	59.9	4279	6 AX105080	AX105080 Sequence
20	2441	59.8	3934	4 CFA419568	AJ419568 Canis fam
21	2441	59.8	4045	12 AF269224	AF269224 Synthetic
22	2441	59.8	4186	6 AX108654	AX108654 Sequence
23	2441	59.8	4195	6 AX108656	AX108656 Sequence
24	2441	59.8	4279	6 AX105082	AX105082 Sequence
25	2438	59.8	4317	4 AF045016	AF045016 Canis fam
26	2438	59.8	4317	6 AX105059	AX105059 Sequence
27	2438	59.8	4323	10 AF286167	AF286167 Rattus no
28	2437	59.7	2752	5 AF099732	AF099732 Fundulus
29	2436	59.7	3840	6 AX481416	AX481416 Sequence
30	2436	59.7	3860	6 AX322787	AX322787 Sequence
31	2436	59.7	3860	6 AX322789	AX322789 Sequence
32	2436	59.7	3988	6 AX024454	AX024454 Sequence
33	2436	59.7	4192	9 AF016535	AF016535 Homo sapi
34	2436	59.7	4264	6 AR051647	AR051647 Sequence
35	2436	59.7	4264	6 AR051650	AR051650 Sequence
36	2436	59.7	4646	6 AX336420	AX336420 Sequence
37	2436	59.7	4646	6 AX336708	AX336708 Sequence
38	2436	59.7	4646	6 AX3391099	AX3391099 Sequence
39	2436	59.7	4646	6 I49610	I49610 Sequence 2
40	2436	59.7	4646	9 HUMMDR1	M14758 Homo sapien
41	2436	59.7	4669	6 I08557	I08557 Sequence 3
42	2436	59.7	6505	6 AR028671	AR028671 Sequence
43	2436	59.7	8630	6 AX012320	AX012320 Sequence
44	2436	59.7	8630	6 AX012321	AX012321 Sequence
45	2436	59.7	9318	6 AR028672	AR028672 Sequence

ALIGNMENTS

RESULT 1

AX339028 AX339028 2856 bp DNA linear PAT 09-JAN-2002
LOCUS Sequence 10 from Patent WO0194400.
DEFINITION AX339028
ACCESSION AX339028
VERSION AX339028.1 GI:18129120
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE Frank, M.H. and Sayegh, M.H.
AUTHORS
TITLE A gene encoding a multidrug resistance human p-glycoprotein
homologue on chromosome 7p15-21 and uses thereof
JOURNAL Patent: WO 0194400-A 10 13-DEC-2001.
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES
source
1..2856
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 888 a 540 c 652 g 776 t
ORIGIN
Alignment Scores:
Pred. No.: 1.87e-296 Length: 2856
Score: 4079.00 Matches: 812
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-873-409-2 (1-812) x AX339028 (1-2856)
Qy 1 MetValAspGluAsnAspIleArgAlaLeuValAsnValArgHisTyrArgAspHisIleGly 20
Db 418 ATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTGGA 477
Qy 21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
Db 478 GTGGTGTAGTCAGAGCGCTGTTTGTTCGGACCAACATCAGTAAATATCAAGTATGGA 537
Qy 41 ArgAspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
Db 538 CGAGATGATGTGACTGATGAAGAGATGGAGAGAGAGAGCAAGCAAGCAAAATCGGTATGAT 597
Qy 61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80
Db 598 TTTATCATGGAGTTTCCTAATAAATTTAATACATTTGGTAGGGGNAAGGAGCTCAAAATG 657
Qy 81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100
Db 658 ATGGAGGCGCAGAAACAGAGGATCGCAATTCGTGCTTGTAGTTGCAAAACCCCAAGATT 717
Qy 101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120
Db 718 CTGATTTTAGATGAGCTACGCTGCTGCTGGATTTCAGAAAGCAAGTCAGCTGTTCAAGCT 777
Qy 121 AlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThr 140
Db 778 GCACGTGGAAGCGCAGCAAGGCTCGGACTACAACTCGTGGTAGCACACCGACTTTCTACT 837
Qy 141 IleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAla 160
Db 838 ATTCGAAGTCAGATTGNTGTGACCTTAAGGATGGNATGCTCGCGGAGAAAGAGCA 897
Qy 161 HisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIle 180
Db 898 CATGCTGAACATAATGCCAAAACGAGCTCTATATTACCTTGTGATGTCACAGGATATT 957
Qy 181 LysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSer 200
Db 958 AAAAAGCTGATGAACAGATGAGTCAATGACATATTCTACTGAAAGAAAGCAACTCA 1017

Qy 201 LeuProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSer 220
Db 1018 CTTCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTCATTGACAAGGCTCAGGAATCC 1077
Qy 221 ThrGlnSerLysGluLysSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 240
Db 1078 ACCCAATCTAAGAGATAGTCTTCTCCTGAAGTCTCTCTATTAAAAAATTTTAAAGTTAAAC 1137
Qy 241 LysProGluTyrProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 260
Db 1138 AAGCCTGAATGGCCTTTTGTGTTCTGGGACATTTGGTCTCTCTAATATGGAATCTGTT 1197
Qy 261 HisProValPheSerIlePheAlaLysIleIleThrMetPheGlyAsnAsnLys 280
Db 1198 CATCAGATTATTTCCATCATCTTTGCCAAAATTTATAACCATGTTTGGAAATAATGATAAA 1257
Qy 281 ThrThrLeuLysHisAspAlaGluLysTyrSerMetIlePheValIleLeuGlyValIle 300
Db 1258 ACCACATTTAAAGCATGATGCAGAAAATTTATTCCCATGATATTCGTCAATTTGGGTGTATT 1317
Qy 301 CysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThr 320
Db 1318 TGCCTTGTCAAGTATTTCATGAGGATTTTACGGCAGACAGGGGAAAATTTTAAAG 1377
Qy 321 MetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAsp 340
Db 1378 ATGAGATTAAAGACACTTGGCCTTCAAAGCCATGTTATATCAGGATATTGCTGTTTGCAT 1437
Qy 341 GluLysGluAsnSerThrGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIle 360
Db 1438 GAAAAGGAAAAACAGCACAGAGGCTTGACAACTATTTAGCCATAGATATAGCACAAAT 1497
Qy 361 GlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeu 380
Db 1498 CAGAGACAAACAGGTTCCAGGATTTGGCGTCTTAACACAAAATGCACATACATGGGACTT 1557
Qy 381 SerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAla 400
Db 1558 TCAGTTATCATTTCTTTATATATGATGGATGGAGATGACATTCCTGATTTCTGATTTGCT 1617
Qy 401 ProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLys 420
Db 1618 CCAGTACTTGGCTGACAGGAATGATTTGAAACCCGACGAATGACTGGGATTTGCCAACAAA 1677
Qy 421 AspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArg 440
Db 1678 GATAGCAAGAACCTTAAGCATGCTGGAAGATAGCACTGAAGCTTTGGAGAAATATAGCT 1737
Qy 441 ThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGln 460
Db 1738 ACTATAGTGTCTATTAAACAAGGGAAGAAAGCCCTTCGAGCAAAATGTATGAAGAGATGCTT 1797
Qy 461 ThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSer 480
Db 1798 ACTCAACACAGAAAATACCTCGAAGAAAACACAGATTTATTGGAAGCTGTTATGCTTAC 1857
Qy 481 HisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGln 500
Db 1858 CATGCCITTTATATTATTTGGCTTATGACAGGCTTCGATTTGGAGCCTATTAAATCAA 1917
Qy 501 AlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAla 520
Db 1918 GCTGGAGCAATGATGCCACAGAGGCGCATGTTCTATAGTTTTTACTGCAATTTGCATATG 1977
Qy 521 MetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAla 540
Db 1978 ATGGCATCGGAAAAACGCTCGTTTGGCTCTCTGAATATTCCAAAGCCAAATCGGGGCT 2037
Qy 541 AlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGly 560
Db 2038 CGGCATCTGTTTGGCTTGTGGAAAAAGAAACCAATATATAGACGCCGCGTCAAGAGGG 2097
Qy 561 LysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrPro 580

Db 2098 AAAAAGCCAGACACATGTAAGGGAAATTTAGAGTTTCGAGAAGTCTCTTTCTCTATCCA 2157
Qy 581 CysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThr 600
Db 2158 TGTGCGCCAGATGTTTTCATCTCGTGGCTATCTCCCTCAGTATTAGCGGAGGAAGACA 2217
Qy 601 VallalaPheValGlySerSerGlyCysGlyLysThrSerValGlnLeuLeuGlnArg 620
Db 2218 GTAGCATTTGTGGGAGCAGCGCTGTGGGAAAAGCACTTCTGTTCAACTTCGACAGAGA 2277
Qy 621 LeuTyrAspProValGlnGlnValLeuPheAspGlyValAlaAlaGluLeuLeuLeu 640
Db 2278 CTTTATGACCCCGTCAGAGCAAGATGCTGTTTGAATGGTGGATGCAAAAGAAATTGAAT 2337
Qy 641 ValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCys 660
Db 2338 GTACAGTGGCTCCGTTCCCAATAGCAATCGTTCTCAAGAGCGCTGTCTCTTCAACTGC 2397
Qy 661 SerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIle 680
Db 2398 AGCATTTGCTGAGAACATCGCTATGCTGACACACAGCGGTGTGGTCCCATTTAGATGAGATC 2457
Qy 681 LysGluAlaAlaAsnAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyr 700
Db 2458 AAAGAAGCCGCAAAATGCAGCAAAATATCATCTCTTTATTTGAAGGTCTCCCTGAGAAATAC 2517
Qy 701 AsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAla 720
Db 2518 AACACACAAAGTTGACTGAAAGGACACAGCTTCTGCGGCGCCAGAAACAAAGACTAGCT 2577
Qy 721 IleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAla 740
Db 2578 ATTGCAAGGGCTCTCTCAAAAACCCAAAATTTTATTGTTGGATGAGGCCACTTCAGCC 2637
Qy 741 LeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArg 760
Db 2638 CTCGATAATGACAGTGAAGGTGTTGAGCATGCGCTTGTATTAAGCCAGGACGGGAAGG 2697
Qy 761 ThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValVal 780
Db 2698 ACATGCTTAGTGGTCACTCACAGGCTCTCGCAATTCAGACGCGAGATTTGNTAGTGGTT 2757
Qy 781 LeuHisAsnGlyLysIleGlyGlnGlnGlyThrHisGlnGlnLeuLeuArgAsnArgAsp 800
Db 2758 CTGCACAAATGGAAGATAAAGAAACAAGAACTCATCAAGAGCTCTCAGAAATCGAGAC 2817
Qy 801 IleTyrPheLysLeuValAsnAlaGlnSerValGln 812
Db 2818 ATATATTTTAAGTTAGTGAATGCAGTCAGTGCAG 2853

RESULT 2
AX339030 AX339030 3177 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 12 from Patent WO0194400.
ACCESSION AX339030
VERSION AX339030.1 GI:18129122
KEYWORDS
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
AUTHORS Frank, M.H. and Sayegh, M.H.

TITLE A gene encoding a multidrug resistance human p-glycoprotein
homologue on chromosome 7p15-21 and uses thereof

JOURNAL Patent: WO 0194400-A 12 13-DEC-2001;

THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)

FEATURES Location/Qualifiers

source 1..3177

/organism="Homo sapiens"

/db_xref="taxon:9606"

BASE COUNT 970 a 601 c 735 g 870 t 1 others

ORIGIN
Alignment Scores:
Pred. No.: 2,16e-296 Length: 3177
Score: 4079.00 Matches: 812
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-873-409-2 (1-812) x AX339030 (1-3177)

Qy 1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
Db 739 ATGTGGATGAGAAATGACATCAGAGCTTTAAATGTGCGGCATTATTCGAGACCATATGGA 798
Qy 21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
Db 799 GTGGTTAGTCAAGAGCCCTGTTTTCGGGACCACCATCAGTAAACATATCAAGTATGGA 858
Qy 41 ArgAspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
Db 859 CGAGATGATGTGACTGATGAAGAGATCGAGAGCAGCAGGAAAGCAAAATGCGTATGAT 918
Qy 61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80
Db 919 TTTATCATGGAGTTTCTTAATAATTTAATACATTTGGTAGGGGAAAAAGGAGCTCAAAATG 978
Qy 81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100
Db 979 AGTGGAGGCGAAGAACAGAGGATCGCAATTCCTGTCCTTAGTTCGAAACCCCAAGATT 1038
Qy 101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120
Db 1039 CTGATTTTAGATGAGGTACGCTCTGCCCTGGATTTCAGAAAGCAAGTCACTGTTCAAGCT 1098
Qy 121 AlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThr 140
Db 1099 GCACCTGGAGAAGGCGAGCAAAAGGTGGGACTACAATCTGTTGGTAGCACACCGGACTTTCTACT 1158
Qy 141 IleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAla 160
Db 1159 ATTCAAGTCAGAGATTGATTTGACCCCTAAAGGATGGAATGCTGCGGAGAAAGAGACA 1218
Qy 161 HisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIle 180
Db 1219 CATGCTGAACTAATGGCAAAACGAGGCTCTATATTATTCACCTTGATGTCACAGGATATT 1278
Qy 181 LysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSer 200
Db 1279 AAAAAAGCTGATGAACAGATGGAGTCAATGACATATTCTACTGAAAGAAAGAACCACTCA 1338
Qy 201 LeuProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluLysSer 220
Db 1339 CTTCTCTCGCACTCTGTGAAGAGCATCAAGTCAGACTTCATTTCAGACGCTGAGGAATCC 1398
Qy 221 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 240
Db 1399 ACCCAATCTAAAGAGATGAAGTCTTCCTGGAAGTCTCTCTATTAAAAATTTTAAAGTTAAAC 1458
Qy 241 LysProGluTrpProPhePheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 260
Db 1459 AAGCTGAATGGCCTTTTGTGGTTCGCGGACATTTGGCTTCTGTTCTTAATGGAACCTGTT 1518
Qy 261 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLys 280
Db 1519 CATCCAGTATTTTCCATCATCTTTTCGAAAAATTTAATCAACATGTTTGGAAATTAATGATATA 1578
Qy 281 ThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIle 300
Db 1579 ACCCATTTAAAGCATGATGCAGAAATTTATTTCATGATATTTCGTCATTTTGGGTGTTATT 1638
Qy 301 CysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyLysLeuLeuThr 320

Db 1639 TGCCTTTGTCAGTTATTTTCATGCGAGGATTTATTTACGCGAGCAGGGGAAATTTTAAACG 1698
Qy 321 MetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAsp 340
Db 1699 ATGAGATTAAAGACACTTGGCTTCAAGCCAGTGTATATACAGGATATTCCTGGTTTGTAT 1758
Qy 341 GluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIle 360
Db 1759 GAAAGGAAACACGACAGGAGGCTTGACACAAATATTAGCCATAGATATAGACAAAT 1818
Qy 361 GlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeu 380
Db 1819 CAAGGAGCAACAGGTTCCAGGATTGGCGTCTTAACACAAATGCAACTAACCTGGGACTT 1878
Qy 381 SerValIleIleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAla 400
Db 1879 TCAGTTATCATTTTCCCTTTATATATGATGGAGATGACATTCCTGATTCGATGTTGCT 1938
Qy 401 ProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLys 420
Db 1939 CCAGTACTTGGCTGACAGGAATGATTGAACCGCAGCAATGACTGGATTTGCCAACAA 1998
Qy 421 AspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArg 440
Db 1999 GATAAGCAAGAACTTAAGCATGCTGGAAGATAGCAACTGAAGCTTTGGAGAATATACGT 2058
Qy 441 ThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGln 460
Db 2059 ACTATAGTGTCAATTAAACAGGGGAAAGGCTTCGAGCAAAATGTATGAAGAGATGCTTCAG 2118
Qy 461 ThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSer 480
Db 2119 ACTCAACACAGAAATACCTCGAAGAAAGCACAGATATTGGAGCTGTATGCATTCAGC 2178
Qy 481 HisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGln 500
Db 2179 CATGCCCTTTATATATTTGCCCTATGCAGCAGGCTTCGATTTGGAGCTTATTAATCAA 2238
Qy 501 AlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAla 520
Db 2239 GCTGACAGTAATGACCCAGGCGCATGTTCAATGATTTTACTGCAATTCATATGAGCT 2298
Qy 521 MetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAla 540
Db 2299 ATGGCCATCGGAAACGCTCGTTTGGCTCTCTGATATTCGAAGCCAAATCGGGGCT 2358
Qy 541 AlaHisLeuPheAlaLeuLeuLysLysProAsnIleAspSerArgSerGlnGluGly 560
Db 2359 GCGCATCTGTTTGGCTTGTGGAAAGAAACCAAAATATAGACAGCGCAGTCAAGAAGG 2418
Qy 561 LysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrPro 580
Db 2419 AAAAGCCACAGACACATGTGAAGGAAATTTAGAGTTTCGAGAGTCTCTTTCTCTATCCA 2478
Qy 581 CysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThr 600
Db 2479 TCTCGCCACAGATGTTTTCATCTCCGTGCTTATCCCTCAGTATTGACGAGGAAAGACA 2538
Qy 601 ValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuGluArg 620
Db 2539 GTAGCATTTGTGGGAGCAGCGCTGTGGGAAAGCACTTCTGTTCAACTTCGACAGA 2598
Qy 621 LeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsn 640
Db 2599 CTTTATGACCCGTCAGAGCAAGTGTGTTGATGGTGTGATGCAAAAGAAATTTGAAT 2658
Qy 641 ValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCys 660
Db 2659 GTACAGTGGCTCGGTCCCAATAGCAATCGTTCTCTCAAGACCTGCTCTTCAACTGC 2718
Qy 661 SerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIle 680

Db 2719 AGCATTGCTGAGAACATCGCCTATTGTTGACAAACAGCCGCTGTGTGCCATTAGATGAGATC 2778
Qy 681 LysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyr 700
Db 2779 AAAGAACCGCCAAATATGCAGCAAAATATCCATTTCTTTATTGAAGGTCTCCCTGAGAAATAC 2838
Qy 701 AsnThrGlnValGlyLeuLysGlyValaGlnLeuSerGlyGlnLysGlnArgLeuAla 720
Db 2839 AACACACAAGTTGGACTGAAAGGAGCAGAGCTTTCTGCGGCCAGAAACAAGACTAGCT 2898
Qy 721 IleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAla 740
Db 2899 ATTGAAGGGCTCTTCTCAAAACCCAAAATTTTATTGTTGGATGAGGCCACTTCAGCC 2958
Qy 741 LeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArg 760
Db 2959 CTCGATAATGACAGTGAAGGTGTTGAGATGCCCTTGTAAAGCCAGGACGGGAAGG 3018
Qy 761 ThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValVal 780
Db 3019 ACATGCTTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAACGCAGATTTGATAGTGGTT 3078
Qy 781 LeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAsp 800
Db 3079 CTGCACAATGGAAGATAAGGAACAGAGAACTCATCAAGAGCTCTCGAGAAATCGAGAC 3138
Qy 801 IleTyrPheLysLeuValAsnAlaGlnSerValGln 812
Db 3139 ATATATTTTAAAGTTAGTGAATGCACAGTCAGTGCAG 3174
RESULT 3
LOCUS AX339032 3621 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 14 from Patent WO0194400.
ACCESSION AX339032
VERSION AX339032.1 GI:18129124
KEYWORDS human.
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Frank, M.H. and Sayegh, M.H.
TITLE A gene encoding a multidrug resistance human p-glycoprotein
homologue on chromosome 7p15-21 and uses thereof
JOURNAL Patent: WO 0194400-A 14 13-DEC-2001;
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES
source
1. 3621
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 1081 a 696 c 842 g 1002 t
ORIGIN
Alignment Scores:
Pred. No.: 2,57e-296 Length: 3621
Score: 4079.00 Matches: 812
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-873-409-2 (1-812) x AX339032 (1-3621)
Qy 1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
Db 1183 ATGTGGATGAGAATGACATCAGAGCTTTAAATGTGGGCAATTATCGAGACCATATTGGA 1242
Qy 21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
Db 1243 GTGGTTAGTCAAGAGCTCTTTTGTTCGGGACCACTCAGTAACAATATCAAGTATGGA 1302
Qy 41 ArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60

Db 1303 CGAGATGATGCTGATGAAAGAGATGAGAGAGCAGCAGGAGGAGCAAAATGCGGTATGAT 1362
Qy 61 PheileMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyValaGlnMet 80
Db 1363 TTTATCATGGAGTTTCCCTAAATAATTTAAATACATTTGGTAGGGGAAAAAGGAGCTCAATG 1422
Qy 81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100
Db 1423 AGTGAGGCGCAGAAACAGAGGATCGCAATTGCTGCTGCTTAGTTTCGAAACCCCAAGATT 1482
Qy 101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValaGlnAla 120
Db 1483 CTGATTTTAGATGAGCTACGCTCGCTCCCTGGATTTCAGAAAGCAAGTCAGCTGTTCAAGCT 1542
Qy 121 AlaLeuGluLysAlaSerLysGlyValArgThrIleValAlaAlaHisArgLeuSerThr 140
Db 1543 GCATCGGAAGAGCGCAGCAAGGTCCGACTACATCTGTTGGTAGCACCCGACTTTCTACT 1602
Qy 141 IleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAla 160
Db 1603 ATTCGAAGTCAGATTGATTGTGACCTTAAGGATGGAATGCTGCGCGAGAAAGAGCA 1662
Qy 161 HisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIle 180
Db 1663 CATGCTGAACATAATGGCAAAACGAGTCTATATTATTCATTGTCATGTCACAGGATATT 1722
Qy 181 LysLysAlaAspGluGlnMetCysSerMetThrTyrSerThrGluArgGlyThrAsnSer 200
Db 1723 AAAAAGCTGATGAACACAGATGGAGTCAATGACATATTTCTACTGAAAGAAAGACCAACTCA 1782
Qy 201 LeuProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluLysSer 220
Db 1783 CTTCCTCTGCACCTCTGTGAAGACATCAAGTCAGACTTCATTGACCAAGCTGAGGAATCC 1842
Qy 221 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLysIleLeuLysLeuAsn 240
Db 1843 ACCCAATCTAAAGAGATAAGTCTTCTGAAAGTCTCTCTATTAAAAAATTTTAAAGTTAAAC 1902
Qy 241 LysProGluTyrProPheValIleValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 260
Db 1903 AAGCTGTAATGGCCCTTTGTGGTCTCTGGGACATTTGGCTTCTGTTCTAAATGGAATGTT 1962
Qy 261 HisProValPheSerIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLys 280
Db 1963 CATCCAGATTTTCCATCATCTTTGCCAAAAATTAACCATGTTTGGAAAAATATGATAAA 2022
Qy 281 ThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIle 300
Db 2023 ACCACATTAAGCATGATGCAGAAATTTATTCATGATATTCGTCAATTTTGGGTGTTATT 2082
Qy 301 CysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThr 320
Db 2083 TGCTTTGTCAAGTATTTCATGAGGATTTATTTTACCGCAGAGCAGGGGAAATTTTAAAG 2142
Qy 321 MetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAsp 340
Db 2143 ATGAGATTAAAGACATTTGGCCTTTCAAGCCATGTTATATACAGATATTGCTGTTGAT 2202
Qy 341 GluLysGluAsnSerThrGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIle 360
Db 2203 GAAAAGGAAAAACAGCACAGGAGGCTTGACAAACATATTAGCCATAGATATAGCACAAAT 2262
Qy 361 GlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeu 380
Db 2263 CAAGGAGCAACAGGTTCCAGGATTTGGCGTCTTAAACACAAATATGCAACTAACATGGGACTT 2322
Qy 381 SerValIleIleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAla 400
Db 2323 TCAGTTATCATTTCTTTATATATGATGAGGAGATGACATTTCTGATTTCTGAGTATTGCT 2382
Qy 401 ProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLys 420

Db 2383 CCAGTACTTGCCGTGACAGGAATGATTGAAACCCAGCAATGACTGGATTTGCCAACAAA 2442
Qy 421 AspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArg 440
Db 2443 GATAAGCAAGAACTTTAGCATGCTTGAAAGATAGCAACTGAAGCTTTGGAGAAATATACGT 2502
Qy 441 ThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGln 460
Db 2503 ACTATAGTGTCTATTAAACAAGGAAAAAGCCCTTCGAGCAAAATGTATGAAGAGATGCTTCAG 2562
Qy 461 ThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSer 480
Db 2563 ACTCAACACAGAAATACCTCGAAGAAAGACAGATATTGGAAGCTGTATGCATTACG 2622
Qy 481 HisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGln 500
Db 2623 CATGCCITTTATATTGCTCTATGCGACGAGGTTTCGATTTGGAGCCTATTTAATTCAA 2682
Qy 501 AlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAla 520
Db 2683 GCTGACCAATGACCCACAGAGGCGATGTTTCATAGTTTTTACTGCAATTTGCATATGGAGCT 2742
Qy 521 MetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAla 540
Db 2743 ATGGCCATCGGAAAAACGCTCGTTTGGCTCTCTGAATATTCAAAGCCAAATCGGGGCT 2802
Qy 541 AlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGly 560
Db 2803 CGCATCTGTTGCTTGTGTTGAAAGAAACCAATATAGACAGCCGCGAGTCAGAGGG 2862
Qy 561 LysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrPro 580
Db 2863 AAAAAGCCAGACACATGTGAAGGAAATTTAGAGTTTCGAGAAGTCTCTTTCTTCTATCCA 2922
Qy 581 CysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThr 600
Db 2923 TGTGCCCCAGATGTTTTCATCTCCGTGGCTTATCCCTCAGTATTGACGAGGAAAGACA 2982
Qy 601 ValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArg 620
Db 2983 GTAGCATTTGTTGGGAGCAGCGCTGTGGGAAAGCACTTCTGTTCACTTCTGCAGAGA 3042
Qy 621 LeuTyrAspProValGlnGlyValLeuPheAspGlyValAspAlaLysGluLeuAsn 640
Db 3043 CTTTATGATACCCCGTCAGAGCAAGTGTCTTTGATGTGTGGATGCAAAAAGAAATGAAT 3102
Qy 641 ValGlnTyrLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCys 660
Db 3103 GTACAGTGGCTCCGCTCCCAATAGCAATCGTTCCTCAAGAGCCTGTGCTCTTCAACTGC 3162
Qy 661 SerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValProLeuAspGluIle 680
Db 3163 AGCATTTGCTGAGACATCGCTATGTTGACAAACAGCCGTGTGGTGCCTATGATGAGATC 3222
Qy 681 LysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGlyLeuProGluLysTyr 700
Db 3223 AAAGAAGCGCAAAATGCAGCAAAATATCCATTTCTTTTATGAAGGTCTCCCTGAGAAATAC 3282
Qy 701 AsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAla 720
Db 3283 AACACAAAGTTGGACTGAAAGGAGCACAGCTTCTGCGCGCCAGAAACAAAGACTAGCT 3342
Qy 721 IleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAla 740
Db 3343 ATTGCNAGGCTCTTCTCCAAAAACCCCAAAATTTTATTTGGATGAGGCCACTTCAGCC 3402
Qy 741 LeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArg 760
Db 3403 CTCGATATGACAGTGAGAGGTGTTTCAGCATGCCCTTTGATAAGCCAGGACGGAGG 3462
Qy 761 ThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValVal 780
Db 3463 ACATGCCCTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAACGCGAGATTTGTATGTTT 3522

QY 781 LeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAsp 800
 |||||
 Db 3523 CTGCAAAATGGAAGATAAAGGAACAAAGGAACATCAAGAGCTCTCTGAAATCGAGAC 3582
 |||||
 QY 801 IleTyrPheLysLeuValAsnAlaGlnSerValGln 812
 |||||
 Db 3583 ATATATTTTAAGTTAGTGAATGCACAGTCAGTGCAG 3618
 |||||

RESULT 4
 LOCUS AX339031 3702 bp DNA linear PAT 09-JAN-2002
 DEFINITION Sequence 13 from Patent WO0194400.
 ACCESSION AX339031
 VERSION AX339031.1 GI:18129123
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1
 Frank, M.H. and Sayegh, M.H.
 A gene encoding a multidrug resistance human p-glycoprotein
 homologue on chromosome 7p15-21 and uses thereof
 Patent: WO 0194400-A 13 13-DEC-2001;
 JOURNAL THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
 FEATURES
 source Location/Qualifiers
 1..3702
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 1098 a 708 c 856 g 1039 t 1 others
 ORIGIN

Alignment Scores:
 Pred. No.: 2,65e-296 Length: 3702
 Score: 4079.00 Matches: 812
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-09-873-409-2 (1-812) x AX339031 (1-3702)

QY 1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
 |||||
 Db 1264 ATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATATTCGAGACCATATTGGA 1323
 |||||
 QY 21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
 |||||
 Db 1324 GTGGTTAGTCAGAGCTGTTTGTTCGGGACCACTCAGTAACATATCAAGTATGGA 1383
 |||||
 QY 41 ArgAspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
 |||||
 Db 1384 CGAGATGATGTGACTGATCAAGAGATGGAGAGCAGCAGCAAGGAAGCAATCGGTATGAT 1443
 |||||
 QY 61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80
 |||||
 Db 1444 TTATATCATGGAGTTTCTTAATAAAATTTAATACATTTGGTGGGAAAAAGAGAGCTCAAAATG 1503
 |||||
 QY 81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100
 |||||
 Db 1504 AGTGGAGGCAGAAACAGAGGATCGCAATGTCTGCGCTTAGTTGCAAAACCCCAAGATT 1563
 |||||
 QY 101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120
 |||||
 Db 1564 CTGATTTTAGATGAGCTACGCTGCGCTGGATTTCAGAAAGCAAGTCAGCTGTTCAAGCT 1623
 |||||
 QY 121 AlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThr 140
 |||||
 Db 1624 GCACGTGGAAGCGCAGCAAGCTCGGACTACAATCGTGGTAGCACACCGAGCTTTCTACT 1683
 |||||
 QY 141 IleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAla 160
 |||||

Db 1684 ATTCGAAGTGCAGATTTTGATTGTGACCCTAAAGAGTGAATCTCGCGGAGAAAGAGCA 1743
 |||||
 QY 161 HisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIle 180
 |||||
 Db 1744 CATGCTGAACATAATGCAAAACGAGGTCTATATTATTCACTTGTGATGTACAGGATATT 1803
 |||||
 QY 181 LysLysAlaAspGluGlnMetClnSerMetThrTyrSerThrGluArgLysThrAsnSer 200
 |||||
 Db 1804 AAAAAAGCTGATGAACAGATGGAGTCAATGACATATTTCTACTGAAGAAGAAAGCAACTCA 1863
 |||||
 QY 201 LeuProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSer 220
 |||||
 Db 1864 CTTCTCTCGCACTCTGTGAGAGCATCAAGTCAGACTTCATTGACAGGCTGAGGAATCC 1923
 |||||
 QY 221 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 240
 |||||
 Db 1924 ACCCAATCTAAAGACATAAGTCTTCTCTGAAGTCTCTCTATTAAAAATTTTAAAGTTAAAC 1983
 |||||
 QY 241 LysProGluTyrProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 260
 |||||
 Db 1984 AAGCCTGAATGCGCTTTTGTGTTCTGGGACATTTGGCTTCTGTTCTAAATGGAACCTGTT 2043
 |||||
 QY 261 HisProValPheSerIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLys 280
 |||||
 Db 2044 CATCCAGTATTTTCCATCATCTTTGCAAAAATATTAACCATGTTTGGAAATATGATAAA 2103
 |||||
 QY 281 ThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIle 300
 |||||
 Db 2104 ACCACATTAAGCATGATGCAGAAATTTATTCCATGATATTCTGTTGGGTGTTATT 2163
 |||||
 QY 301 CysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThr 320
 |||||
 Db 2164 TGCTTTGTGAGTTATTTCATGAGGGATTATTATACGCGCAGCAGGCGGAAATTTTAAACG 2223
 |||||
 QY 321 MetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAsp 340
 |||||
 Db 2224 ATGAGATTAAAGACACTTGGCCTTCAAGCCATGTTATATCAGGATATTGCTCGTTGAT 2283
 |||||
 QY 341 GluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIle 360
 |||||
 Db 2284 GAAAAGGAAAAACAGCACAGGAGGCTTGACAAACATATTAGCCATAGATATAGCACAAATT 2343
 |||||
 QY 361 GlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetClyLeu 380
 |||||
 Db 2344 CAAGAGCAACAGGTTCCAGGATTTGGCGTCTTAACACAAAATGCAACTAACATGGGACTT 2403
 |||||
 QY 381 SerValIleIleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAla 400
 |||||
 Db 2404 TCAGTTATCATTTCTTTTATATGATGGATGGAGATGACATTCCTGATTCTGAGTATTGCT 2463
 |||||
 QY 401 ProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLys 420
 |||||
 Db 2464 CCAGTACTTGGCGTGACAGGAATGATTGAAACCGCAGCAATGACTGGATTTCGCCAACAA 2523
 |||||
 QY 421 AspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArg 440
 |||||
 Db 2524 GATAAGCAAGAACTTAAGCATGCTGGAAAGATAGCAACTGAAGCTTTTGGAGATATACGT 2583
 |||||
 QY 441 ThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGln 460
 |||||
 Db 2584 ACTATAGTGTCTTAACAAAGGAAAGAACCTTCGAGCAAAATGTTATGAAGATGCTTCAG 2643
 |||||
 QY 461 ThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSer 480
 |||||
 Db 2644 ACTCAACACAGAAATACCTCGAAGAAAGCAGACAGATTATTGGAAGCTGTTATGCAATTCAGC 2703
 |||||
 QY 481 HisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrIleuIleGln 500
 |||||
 Db 2704 CATGCTTTATATATTTTGGCTTATGCGACGAGGTTTCGATTGGAGCGCTATTAAATTCAA 2763
 |||||
 QY 501 AlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAla 520
 |||||
 Db 2764 CTGGACGATGACCCCGACAGGGCATGTTTCATAGTTTTTACTGCATATGCGAGCT 2823
 |||||

Qy 521 MetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAla 540
Db 2824 ATGGCCATCGGAAACACGCTGTTTGGCTCTCTGAATATTCAGAACGCAAAATCGGGGCT 2883
Qy 541 AlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGly 560
Db 2884 GCGCATCTGTTTGGCTTGTGGAAAGAAACCAATATAGACAGCCGAGTCAAGAGGG 2943
Qy 561 LysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrPro 580
Db 2944 AAAAGCCAGACACATGTGAAGGAATTTAGAGTTTCGAGAAGTCTCTTCTCTATCCA 3003
Qy 581 CysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThr 600
Db 3004 TGTGCCCCAGATGTTTTCATCTCCGTGGCTTATCCCTCAGTATTGAGCGAGGAAAGACA 3063
Qy 601 ValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuGlnArg 620
Db 3064 GTAGCATTTGTGGGAGCAGCGGCTGTGGGAAAGACATCTCTGTTCAACTTCTTCAGAGA 3123
Qy 621 LeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsn 640
Db 3124 CTTATGACCCCGTCAAGGACAGTGTCTTTGTATGTGTGGATGCAAGAAATTTGAAT 3183
Qy 641 ValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCys 660
Db 3184 GTACAGTGGCTCCGTTCCCAATACCAATCGTCTCTCAAGAGCCGTGCTCTTCAACTGC 3243
Qy 661 SerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIle 680
Db 3244 AGCATGCTGAGAACATCGCTATGTTGTCACACACGCGGTGTGTGCCATTAGATGAGATC 3303
Qy 681 LysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyr 700
Db 3304 AAAGAGCCGCAATGCGACCAATATCCATCTTTTATTGAAGTCTCCCTGAGAAATAC 3363
Qy 701 AsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAla 720
Db 3364 AACACACAAGTTGGACTGAAGGAGCAGACGCTTTCTGGCGGCCAGAAACAAAGACTAGCT 3423
Qy 721 IleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAla 740
Db 3424 ATTGCAAGGCTCTTCTCCAAAACCCCAAAATTTTATTGTTGGATGAGGCCACTTCAGCC 3483
Qy 741 LeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArg 760
Db 3484 CTCGATAATGACAGTGAGAGGTGTTGAGCATGCCCTTGATAAGCCAGGACGGGAAGG 3543
Qy 761 ThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValVal 780
Db 3544 ACATGCCCTAGTGTCTCCTCACAGGCTCTCTGCAATTCAGAACGCAGATTTGATAGTGT 3603
Qy 781 LeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlnLeuLeuArgAsnArgAsp 800
Db 3604 CTGCACAAATGGAAGATAAAGAAACAAGAACTCATCAAGAGCTCTCTGAGAAATCGAGAC 3663
Qy 801 IleTyrPheLysLeuValAsnAlaGlnSerValGln 812
Db 3664 ATATATTTAAGTTAGTGAATGCAGTCAGTGCAG 3699
RESULT 5
LOCUS AX478104 3699 bp DNA linear PAT 12-AUG-2002
DEFINITION Sequence 34 from Patent WO0240541.
ACCESSION AX478104
VERSION AX478104.1 GI:22217064
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1

AUTHORS Tang, Y. T., Yue, H., Nguyen, D. B., Hafalia, A. J., Elliott, V. S., Lu, Y.,
Walia, N. K., Yao, M. G., Baughn, M. R., Gandhi, A. R., Ding, L.,
Sanjanwala, M., Ramkumar, J., Arvizu, C., Gietzen, K. J., Lal, P. G.,
Azimzai, Y., Khan, F. A., Thangavelu, K., Thornton, M., Lu, D. A.,
Tribouley, C. M., Warren, B. A., Ison, C. H., Das, D., Raumann, B. E.,
Policky, J. L. and Kearney, L.
TITLE Transporters and ion channels
JOURNAL Patent: WO 0240541-A 34 23-MAY-2002;
Incyte Genomics, Inc. (US)
FEATURES
source 1. 3699
Location/Qualifiers /organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 1116 a 707 c 860 g 1016 t
ORIGIN
Alignment Scores: 2.62e-284 Length: 3699
Pred. No.: 3919.00 Matches: 785
Score: 96.80% Conservative: 1
Percent Similarity: 96.67% Mismatches: 0
Best Local Similarity: 96.67% Indels: 26
Query Match: 96.08% Gaps: 1
DB: 6
US-09-873-409-2 (1-812) x AX478104 (1-3699)
Qy 1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
Db 1339 ATGTTGGATGAGATGATCATCGAGCTTTAAATGTGGCATTTATCGAGACCATATTGGA 1398
Qy 21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
Db 1399 GTGGTTAGTCAAGAGCCTGTTTTCGGGACCACCATCAGTAACAATATCAAGTATGGA 1458
Qy 41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
Db 1459 CGAGATGATGTGACTGATGAAGAGATGGAGAGCAGCAGCGGAAGCAAAATGCGGTATGAT 1518
Qy 61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyValaGlnMet 80
Db 1519 TTTATCATGGAGTTTCTTAATAATTTAATACATTTGTTAGGGGAAAAAGGAGCTCAATG 1578
Qy 81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100
Db 1579 AGTGAGGCGCAGAAACAGAGGATCGCAATTGCTGTCCTTAGTTCGAAACCCCAAGATT 1638
Qy 101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120
Db 1639 CTGATTTTAGATGAGGCTACGCTCTGCCCTGGATTTCAGAAAGCAAGTCAAGCTGTTCAAGCT 1698
Qy 121 AlaLeuGluLysAlaSerLysGlyValArgThrThrIleValValAlaHisArgLeuSerThr 140
Db 1699 GCATGGGAAGGCGAGCAAGGTCGGACTCAATCTGTGGTAGCACACCGACTTTCTACT 1758
Qy 141 IleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAla 160
Db 1759 ATTGCAAGTCAGATTGATTGTGACCTTAAAGGATGGAATGCTGGCGGAGAAAGAGCG 1818
Qy 161 HisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIle 180
Db 1819 CATGCTGAACATAATGGCAAAACGAGGCTATATTATTACATTGTGTGATGTCAGGATATT 1878
Qy 181 LysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSer 200
Db 1879 AAAAAGCTGATGAACAGATGAGTCAATGACATATTCTACTGAAAGAAAGACCAACTCA 1938
Qy 201 LeuProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSer 220
Db 1939 CTTCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTCATTGACAAAGCTGAGGAATCC 1998
Qy 221 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 240

Db 1999 ACCCAATCTAAAGAGATAAGTCTCTCTGAAGTCTCTCTATTAAAAATTTTAAAGTTAAAC 2058
Qy 241 LysProGluTyrProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 260
Db 2059 AAGCCGTGAATGACCTTTTGGTCTGGGGACATTTGGCTCTCTTAATGGAATGTT 2118
Qy 261 HisProValPheSerIlePheAlaTyrIleThrMetPheGlyAsnAsnAspLys 280
Db 2119 CATCCAGTATTTTCCATCATCTTTGCAAAATTTATAACCATGTTTGGAAATTAATGATAA 2178
Qy 281 ThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIle 300
Db 2179 ACCACATTAAGCATGATGACGAAATTTTATCCATGATATTCGTATTTGGGTGTTAT 2238
Qy 301 CysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThr 320
Db 2239 TCGTTTGTCAAGTATTTTCATG----- 2259
Qy 321 MetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAsp 340
Db 2260 -----CAGGATATTGCTTGGTTTGAT 2280
Qy 341 GluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIle 360
Db 2281 GAAAAGGAAAACAGCACAGAGGCTTGAACAATATTAGCCATAGATATAGCAAAAT 2340
Qy 361 GlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeu 380
Db 2341 CAAGGAGCAACAGTTCCAGGATTTGGCGTCTTAACACAATAATGCAACTAACATGGGACTT 2400
Qy 381 SerValIleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAla 400
Db 2401 TCAGTTATCATTTTCTTTATATATGATGCGAGATGACATTCCTGATTCAGTATTGCT 2460
Qy 401 ProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLys 420
Db 2461 CCAGTACTTTCGCGTACAGGAATGATTTGAACCGCAGCAATGACTGGATTTGCCAACAAA 2520
Qy 421 AspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArg 440
Db 2521 GATAGCAGAACTTAAGCATGCTGGAAAGATAGCACTGAAGCTTTGGAGATATACGT 2580
Qy 441 ThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGln 460
Db 2581 ACTATAGTGTCTATTAACAAGGGAAAAGCCTTCGAGCAAAATGTATGAGAGATGCTTCAG 2640
Qy 461 ThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSer 480
Db 2641 ACTCAACACAGAAATACCTCGAAGAAAGCACAGATTTATTTGGAGCTGTTATGCATTCAGC 2700
Qy 481 HisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGln 500
Db 2701 CATGCTTTATATATTTGCTTATGCGGAGGGTTTCGATTTGGAGCCTATTATTCNA 2760
Qy 501 AlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAla 520
Db 2761 GCTGACGAATACCCACAGAGGCGATGTTTCATAGTTTTTACTGCAATTCATATGAGCT 2820
Qy 521 MetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAla 540
Db 2821 ATGGCCATCGGAACACGCTCGTTTGGCTCTCGAATATTCAAAGCCAAATTCGGGGCT 2880
Qy 541 AlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGly 560
Db 2881 GCGCATCTGTTTGGCTTTGTTGAAAAGAAACCAATATATAGACAGCGCGAGTCAGAGGG 2940
Qy 561 LysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrPro 580
Db 2941 AAAAGCCAGACACATGTCAAGGAAATTTAGAGTTTCGAGAGTCTCTTTCTTCATCCA 3000
Qy 581 CysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThr 600
Db 3001 TGTGCGCCAGATGTTTTCATCTCTCGTGGCTTATCCCTCAGTATTGAGCGAGGAAGACA 3060

Qy 601 ValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArg 620
Db 3061 GTAGCATTTTGGGGAGCAGCGCTGTGGAAAAAGCACTTCTGTTCACACTTCTGCAGAGA 3120
Qy 621 LeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsn 640
Db 3121 CTTTATGACCCCGTCAAGGACAAGTGTCTGTGTATGTGTGTGATGCAAAAGAATTGAAT 3180
Qy 641 ValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCys 660
Db 3181 GTACAGTGGCTCCGTTCCCAATAGCAATCGTTCTCAAGAGCCTGTGCTCTTCAACTGC 3240
Qy 661 SerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValProLeuAspGluIle 680
Db 3241 AGCATTGCTGAGAACATCGCTATGTCACAACAGCCGTGTGGTGCCATTAGATGAGATC 3300
Qy 681 LysGluAlaIleAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyr 700
Db 3301 AAAGAAGCCGCAAAATGCAGCAAAATATCCATTTCTTTTATTGAAGGTCTCCCTGAGAAATAC 3360
Qy 701 AsnThrGlnValGlyLeuLysGlyValAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAla 720
Db 3361 AACACACAAGTTGGACTGGAAGGACACAGCTTCTGCGCGCCAGAAACAAGACTAGCT 3420
Qy 721 IleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAla 740
Db 3421 ATTTCAAGGGCTCTTCTCCAAAACCCCAAAATTTTATTGTTGGATGAGGCCACTTCAGCC 3480
Qy 741 LeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArg 760
Db 3481 CTCGATAATGACAGTGAAGGTGTTTCAGCATGCGCTTGTATAAAGCCAGGACGGGAAGG 3540
Qy 761 ThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValVal 780
Db 3541 ACATCCCTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAACGCGAGATTTGATAGTGT 3600
Qy 781 LeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAsp 800
Db 3601 CTGCACATGAAAGATAAAGGAACAAGAACTCATCAAGAGCTCTCAGAAATCAGAC 3660
Qy 801 IleTyrPheLysLeuValAsnAlaGlnSerValGln 812
Db 3661 ATATATTTTAAAGTTAGTGAATGCACAGTCAGTGCAG 3696

RESULT 6
AX339027 AX339027 2066 bp DNA linear PAT 09-JAN-2002
LOCUS Sequence 9 from Patent WO0194400.
DEFINITION AX339027
ACCESSION AX339027
VERSION AX339027.1 GI:18129119
KEYWORDS
SOURCE human.

ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Frank, M.H. and Sayegh, M.H.
TITLE A gene encoding a multidrug resistance human p-glycoprotein
homologue on chromosome 7p15-21 and uses thereof
JOURNAL Patent: WO 0194400-A 9 13-DEC-2001;
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)

FEATURES
source
1. 2066
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 647 a 403 c 470 g 546 t
ORIGIN

Alignment Scores:
Pred. No.: 6.27e-250 Length: 2066
Score: 3456.00 Matches: 687
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 84.73% Indels: 0
DB: 6 Gaps: 0
US-09-873-409-2 (1-812) x AX339027 (1-2066)

QY 126 SerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAsp 145
DB 3 AGCAAGGTCGAGCTACATCGTGGTAGCACACCGACTTCTTACTATTTCGAAAGTCAGAT 62

QY 146 LeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyValAlaHisAlaGluLeuMet 165
DB 63 TTGATTGTGACCTTAAGAGATGGAATGCTGGCGGAGAAAGGAGACATGCTGCACTAATG 122

QY 166 AlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGlu 185
DB 123 GCAAAACGAGGCTATATTATTACTTGTGATGTCACAGGATATTAAAAAGCTGATGAA 182

QY 186 GlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSer 205
DB 183 CAGATGGAGTCAATGACATATTCTACTGAAGAAAGACCAACTCACTTCTCTGCACCTCT 242

QY 206 ValLysSerIleLysSerAspPheIleAspLysAlaGluLysThrGlnSerLysGlu 225
DB 243 GTGAAGAGCATCAAGTCAGACTTCATTGTGAAGGCTGAGGAATCCACCATCTTAAGAG 302

QY 226 IleSerLeuProGluValSerLeuLysIleLeuLysLeuAsnLysProGluTrpPro 245
DB 303 ATAAGCTTCTCGTAGTCTCTATTAAAAATTTAAAGTTAAACAGGCTGAATGGCT 362

QY 246 PheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSer 265
DB 363 TTTGTGGTTCGGGACATGGCTCTGTCTTAATGGAATGTTCAATCGAGTATTTCCTCC 422

QY 266 IleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThrLeuLysHis 285
DB 423 ATCATCTTTGGCAAAATTAATACCATGTTTGGAAATAATGATAAAACCATTAAGCAT 482

QY 286 AspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyr 305
DB 483 GATGCAGAAATTTATCCATGATATTCGTCAATTTGGTGTTATTGCTTTGTCAGTTAT 542

QY 306 PheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHis 325
DB 543 TTCATGCAGGATATTTTACGCGACAGCAGGGGAAATTTTAACGATGAGATTAAGACAC 602

QY 326 LeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSer 345
DB 603 TTTGGCTTCAAAGCCATGTTATATCAGGATATTGGCTGTTTGTGATAAAGGAAACAGC 662

QY 346 ThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGly 365
DB 663 ACAGAGGCTTGACAAACATATTAGCCATAGATATAGCAAAATTCAGGAGCAACAGGT 722

QY 366 SerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSer 385
DB 723 TCCAGATTGGCGCTTAACACAAATGCAACTAACATGGGACTTTCAGTTATCATTTCC 782

QY 386 PheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaVal 405
DB 783 TTTATATATGATGGAGATGATCATTTCTGATTCTGAGTATTGCTCCAGTACTTCCCGTG 842

QY 406 ThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeu 425
DB 843 ACAGGAATATTGAAACCCAGCAATGACTGGATTTTGCCCAAAAGATAAGCAAGAACTT 902

QY 426 LysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeu 445
DB 903 AAGCATGCTGGAAGATACCAACTGAAGCTTTGGAGATATACGTACTATAGTGTCAITA 962

QY 446 ThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsn 465
DB 963 ACNAGGAAAAAGCCTTCGAGCAAAATGTTATGAAGATGCTTTCAGACTCAACACAGAAAT 1022

QY 466 ThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyr 485
DB 1023 ACCTCGAAAGAACACACAGATTATTGGAAGCTGTTATGCAATTCAGCCATGCTTTATATAT 1082

QY 486 PheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThr 505
DB 1083 TTTGGCTATGAGCAGAGGTTTCGATTTTGGAGCCTATTAAATTCAGCTGGAGCAATGACC 1142

QY 506 ProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLys 525
DB 1143 CCAGAGGCAATGTTTCATAGTTTTCGCAATTCATATGAGCTATGCGCATCGGAAAA 1202

QY 526 ThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAla 545
DB 1203 ACGTCTGTTTGGCTCTCTGAATATTCCAAAGCCAAATCCGGGGCTGCGCATCTGTTTGC 1262

QY 546 LeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThr 565
DB 1263 TGTTTGGAAAAAGAACCAATATAGACAGCCGAGTCAAGAGGGGAAAGCCAGACACA 1322

QY 566 CysGluGlyAsnLeuGluPheArgGluValSerPheTyrProCysArgProAspVal 585
DB 1323 TGTGAAGGAATTTAGAGTTTCGAGAGCTCTTCTTCTATCCATGTCGCCAGATGTT 1382

QY 586 PheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGly 605
DB 1383 TTTCACTCTCGTGGCTTATCTCAGTATTGAGCGAGGAAAGACAGTAGTATGCTTTGGGG 1442

QY 606 SerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProVal 625
DB 1443 AGCAGCGCTGTGGGAAAAAGCACTTCTGTTCAACTTCTGCAGAGACTTTATGACCCCGTG 1502

QY 626 GlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArg 645
DB 1503 CAAGACAAAGTGTGTTTGATGGTGGATGCCAAAGAAATTTGAATGTACAGTGGCTCCGT 1562

QY 646 SerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsn 665
DB 1563 TCCCAATAGCAATCGTTCTCTCAAGAGCCTGTGCTCTTCACTGCAGCATTTGCTGAGAAC 1622

QY 666 IleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsn 685
DB 1623 ATCGCTATGTTGTCACACAGCGCTGTGGTGCCTATTAGATGAGATCAAAAGAGCCGCAAT 1682

QY 686 AlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGly 705
DB 1683 GCAGCAAAATATCCATTTTATTGAAAGTCTTCCCTGAGAAATACACACACAGTTGA 1742

QY 706 LeuLysGlyAlaGlnLeuSerGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeu 725
DB 1743 CTGAAGGAGCACACAGCTTTCTGGCGCCAGAAAAACAAGACTAGCTATTGCAAGGCTCTT 1802

QY 726 LeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSer 745
DB 1803 CTCCAAAAACCCAAAAATTTTATTGTTGGATGAGGCCACTTCAGCCCTCGATAATCACAGT 1862

QY 746 GluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValVal 765
DB 1863 GAGAAAGTGGTTACAGCATGCCCTTGATAAAGCCAGGAGCGGAAGGACATGCCCTAGTGGTC 1922

QY 766 ThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLys 785
DB 1923 ACTCACAGCTCTCTGCAATTCAGAACCGCAGATTGTAGTGGTTCTGCACATCGAAAG 1982

QY 786 IleLysGluGlnGlyThrHisGlnGluLeuArgAsnArgAspIleTyrPheLysLeu 805
DB 1983 ATAAAGGAACAAGGAACCTCATCAAGAGCTCTCGAGAAATCGAGACATATATTTAAGTTA 2042

QY 806 ValAsnAlaGlnSerValGln 812
DB 2043 GTGAATGCACAGTCAAGTCAG 2063

RESULT 7
GGA9799
LOCUS Gallus gallus mRNA for ABC transporter protein. 4018 bp mRNA linear VRT 03-MAR-1999
DEFINITION ABC transporter protein; ccmdr1 gene; P-glycoprotein.
ACCESSION AJ009799
VERSION 1
KEYWORDS ABC transporter protein; ccmdr1 gene; P-glycoprotein.
SOURCE Gallus gallus.
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
1 (bases 1 to 4018)
Edelmann,H.M., Duchek,P., Rosenthal,F.E., Foger,N., Glackin,C., Kane,S.E. and Kuchler,K.
Ccmdr1, a chicken P-glycoprotein, confers multidrug resistance and interacts with estradiol
Biol. Chem. 380 (2), 231-241 (1999)
99209805
10195430
source
location/Qualifiers
1. .4018
/organism="Gallus gallus"
/db_xref="taxon:9031"
/tissue_type="intestine"
109. .3975
/gene="ccmdr1"
109. .3975
/gene="ccmdr1"
/note="ABC transporter protein"
/codon_start=1
/product="P-glycoprotein"
/protein_id="CAA08835.1"
/db_xref="GI:3355757"
/translation="MHSEDEKQKHTVDGNTVEIANSNODQPEDSKKKKKKKPOM
VSPLAFRYSCDKLMLFGSLIAIHGTSPLIAWIIIFGMDTDSFVSGMNTNGS
SGNSADFNKLEEMTRYATYISALAAVLVAAIYQTSFVWLAAGQVKIRKFF
HAIMEQETGFWDNAGELNRLIDVSKINEIGDKGLFIQSELTFLTGFVGFIR
GKWLTLVAVSPVLGAALWAKILTFTDKEQAYAKAGAVEVLSAVRTVIAFG
GQEKLEIKRYHNLDAKRGIRKATISNMGAAFLIYASVALAFWYGTLLILANEY
STGNLTVFFSLVIGAFSIGOTAPSIKAFANARGAAVAFNIIDNEPRIDSVDAHK
PHIKGNLEFQNFVPSRDPVEILKGLNLKNGCQTVLVGSGCGSKSTVQLIQ
FYDPKEGTTIDGDLKLVNLYLREIIGVNVQEPVLFATTIAENIRYRSDVTMEEI
ERATKEANAYDFIMKPKFTVVGEGAQMSGGQKQRIATARALVHNPKILLDEAT
SALDTSESVVQALDKAREGRTTVVVAHRLSTVRNADLIAVFSGVITQGNHSLI
EKKGIYKLVNMQTITETDPSSEKSENAVSVKRSQSNLDLSLKKELRRGSTRSMK
KFGENDDEKSSPDEBLPVVSLKMLKNEMPHYVAGTFCFAIUNGALQAPSVI
FSEIIGITSETQKVRKNSLYSLFLALGIIISFTFVQSFAPGKAGELTUMKRF
MAFKAMLRQDMAFDNDKNGTALTRLANDASQVKGATVGRFLALIAQNLNGTII
ISLVYQWGLTLLLVAVPIIYAVNGMIEMKLAGHAKDKIELEAGKATAEIENRT
VASLRKRELFELLYGHEILLVYRNGVKKXAHIPGFCFSLSQAMFFTYAGCFRFGAYLV
VNGHTEYKTVLRFVSAVFGMALGQTSFADYAKAKISAAHLVFLNRPVPPIDSVR
EDGEKPEFGGNTRIKQVKNYPNRPVKILQGLNLAVKGETLALVSGSGGSKSTV
QLLERFYPLSGEIRVDDIDAKTLNQLMRSHIGVSRPILFDFTIAENIAYGNRSR
EVSHEEIIISAAKAAIHGFIIDSLPEKYNTRVGDGKQLSGQKQRIATARALIRKPI
LILLDEKTSALDTESEKIVQEALDKAREGRTCTIVIAHRLSTIQNADKIAVIQNGKVIQ
GTHQQLAEKGFYSLVNVQSGCNM"
112. .3972
mat_peptide
/gene="ccmdr1"
/product="ABC transporter protein"
/note="P-glycoprotein"
BASE COUNT 1288 a 758 c 911 g 1061 t
ORIGIN
Alignment Scores:
Pred. No.: 2,32e-176 Length: 4018

Score: 2480.00 Matches: 472
Percent Similarity: 75.70% Conservative: 151
Best Local Similarity: 57.35% Mismatches: 186
Query Match: 60.80% Indels: 14
DB: 5 Gaps: 1
US-09-873-409-2 (1-812) x GGA9799 (1-4018)
QY 2 ValAspGluAsnAspIleAArgAlaLeuAsnValAArgHisTyrAArgAspHisIleGlyVal 21
Db 1489 ATTGACGGCAGGATCTGAAGAGCCTGAATGTAGATATCTCGAGAGATTATTGGTGTG 1548
QY 22 ValSerGlnGluProValIleuPheGlyThrIleSerAsnAsnIleIleYsTyrGlyArg 41
Db 1549 GTGAACACAGGAGCCGCTGCTTCTCACTATTTGCAAAAATATTGCTTATGCCCT 1608
QY 42 AspAspValThrAspGluGluMetGluArgAlaAaArgGluAlaAsnAlaTyrAspPhe 61
Db 1609 GAGGATGTCACCATCGAAGAAATTCGAAGAGCTACCAAGGAGCTAATGCTTATGTTT 1668
QY 62 IleMetGluPheProAsnLysPheAsnThrIleuValGlyGluLysGlyAlaGlnMetSer 81
Db 1669 ATCATGAAGCTCCCAAGAAGTTTGAACCTGTGTGGTGGAGAAAGAGGGGCACAGATG 1728
QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProIleIleu 101
Db 1729 CGAGGGCAGAGCAGCAATAGCAATTCGCTGCTGTCTCACAATCTCTAAATTTCT 1788
QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
Db 1789 CTGCTTGTATGAGCAACGTCAGCTTGTGATCTAGAGATGAATCGTGTGGAGGAGCA 1848
QY 122 LeuGluLysAlaSerLysGlyArgThrIleValValAlaAlaHisArgLeuSerThrIle 141
Db 1849 CTGGACAAGGCAAGAGAGGACGACCGTGTGTAGTAGCTCATCGTCTATCAACAGTC 1908
QY 142 ArgSerAlaAspLeuIleValThrIleuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
Db 1909 CCAATGTCAGATCTTATCGCTGTGTGAAAGTGGAGTTATCACAGAACAAGAAATCAT 1968
QY 162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIleLys 181
Db 1969 TCTCACTGATGAAAAAAGGATCTACTATAAATGTTAAACATGCAGACAATAGAA 2028
QY 182 LysAlaAspGluGlnMetGlnSerMetThrTyrSerThrGluArgLysThrAsnSerLeu 201
Db 2029 ACTGAAGATCCATCATCATGAAAAAAGTGAATAATGCAGTCAGTGTCAAAAAGAGTGATCT 2088
QY 202 ProLeuHisSerValLysSerIleLysSerAspPhe----- 213
Db 2089 CAATCAAAATCTGGATGAATCCCTAAAAAAGAAATTTGAGAAGAGGATCAACTAGGAGAAGC 2148
QY 214 -----IleAspLysAlaGluGluSerThrGlnSerLysGluIleSer 227
Db 2149 ATGAAAAAGCCAGGAGAACCTTAATGATCTGATGAGAAAGGTAGTTTACCTGATGAGNA 2208
QY 228 LeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTyrProPheVal 247
Db 2209 TTACCTCCAGTTTTCATTTCTGAAGAACTTATGAAGTTTGAACAAAACGAATGGCCATCTTT 2268
QY 248 ValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIle 267
Db 2269 GTAGCTGGAACCTTCTGTGCAATTTGTAATGGAGCTTTACAAACCTGCATTTTTCAGTCA 2328
QY 268 PheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAla 287
Db 2329 TTCTCAGAAATTTATAGGATTTTTTCAGAAACTGATCAGAAAGTTTTCAGAGAAAAGAGC 2388
QY 288 GluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMet 307
Db 2389 AACTTATATTCACTGCTGTTTTCAGTCTGGATCAATTCCTCTTTTTCATCTTCTTGT 2448
QY 308 GlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuAla 327


```

Db 2449 CAGGGTTTTCATTTGGCAAACTGGAGAAATCTTACAAATGAAGCTTCGATTCAATGCA 2508
Qy 328 PheLysAlaMetLeuTyrGlnAspIleAlaTrrPheAspGluGlyGluAsnSerThrGly 347
Db 2509 TTTAAAGCAATCTGAGACAGACATGGCTGTTTGTATGATCTCTAAATAATAGCACTGA 2568
Qy 348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg 367
Db 2569 GCATGACAAACAAGACTTGTCTAATGATGCTCTCACAAAGTAAAGAGGACCACTGGTGTCA 2628
Qy 368 IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387
Db 2629 TTGGCAATTAATTTGCCAAACATAGCTAACCTGGGAACCTGGAAATATTATATCTTTAGTT 2688
Qy 388 TyrGlyTrrPheLeuMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGly 407
Db 2689 TATGCTGGCAGCTGACCTGTACTCTTAGCTGTGTACCAATCATTCGACTGGCAGGA 2748
Qy 408 MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHis 427
Db 2749 ATGATTGAAATGAAGATGTTGGCTGGCATGCAAAAAGATAAAATAGAGCTGGAAGCT 2808
Qy 428 AlaGlyLysIleAlaThrGluAlaLeuGluAsnIleAtrqThrIleValSerLeuThrArg 447
Db 2809 CGAGGAAGATTTGCCACAGAGCCATAGAAAATATTCGAACTGTTGCTAGTTTGACTCGG 2868
Qy 448 GluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSer 467
Db 2869 GAAAAAGATTTGAGCTCATGTACGGAGAACATTTGCTGTACCATACAGAAATTCGTGTG 2928
Qy 468 LysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487
Db 2929 AAGAAGGGCCACATATTTGGGTTTTGTTTTTCCCTTTCACAAAGCAATGATGTTTTTACC 2988
Qy 488 TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507
Db 2989 TATGCTGGCTGTTTCGCTTTGGAGCTATCTGTGGTTAAATGGACACATAGAGTACAAA 3048
Qy 508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeu 527
Db 3049 ACTGTGTTTTAGTGTTCAGCTGTGTATTGTTGGTGTATGGCATTAGACAAACCAGC 3108
Qy 528 ValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeu 547
Db 3109 TCTTTTGGACCAAGACTATGCCAAAGCCAGATATCAGCAGCCCACTGTTGTTGCTGTT 3168
Qy 548 GluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAsnThrCysGlu 567
Db 3169 AATAGAGTACCTCCAATAGACAGTTTACCCGAGGATGGAGAGAAACCTGAAAAATTTGGG 3228
Qy 568 GlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIle 587
Db 3229 GGAACACAGAGATCAAGATGTGAATTTAACTATCCAAACCGCAGAGGTCAAAATC 3288
Qy 588 LeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSer 607
Db 3289 CTCCAAGGCTTGAATCTAGCAGTAGAAAAAGGAGAAACGTTGGCCCTTGTGTTAGCAGT 3348
Qy 608 GlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGly 627
Db 3349 GGATGTGGAAGAGCAGCTGTTGTCAGCTACTTGAGAGATTTTATGATCCACTCAGTGGA 3408
Qy 628 GlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrrPheLeuArgSerGln 647
Db 3409 GAAATGTGTTTGTATGACATTTGATGCAAGACACTAAATATTCAGTGGCTGAGATCTCAC 3468
Qy 648 IleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAla 667
Db 3469 ATCGGTATGCTCTCAAGAGCCAAATCCTGTTCCGACTTCACCATCGCTGAAAAACATCGG 3528
Qy 668 TyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaIle 687

```

```

Db 3529 TACGGGGNCAACAGCTCGGAGGTGTCCTCCATCAGGAAATCATTAGTCAGCGCAAAGCAGCC 3588
Qy 688 AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLys 707
Db 3589 AGTATTATTCTTTCATTGACTCTCTGCCAGAAAAATACAATACTCGTGTAGGAGACAAG 3648
Qy 708 GlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGln 727
Db 3649 GGAACACAGCTTTCTGGTGGTCAGAAAACAGTATTGCTATTGCTCGAGCTCTTATACGT 3708
Qy 728 LysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 747
Db 3709 AAGCCCCAGATTTCTACTCTGATGAGCTACATCTGCTTGGACACAGAAAGCGAAAAG 3768
Qy 748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767
Db 3769 ATTGTCCAGGAAGCACTGATATAAAGCCAGAGAAAGGTGCGACCTGCATCGTATAGCTCAC 3828
Qy 768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys 787
Db 3829 CGCCTGTCCACCATCCAGAAATGCTGACAAGATTGCTGTGATCCAGAACCGCAAGGTCAAT 3888
Qy 788 GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807
Db 3889 GAGCAGGAGACTCACCAGCAGCTCTGCTGCTGAAAAGGCTTCTACTACTCGTGGTCAAT 3948
Qy 808 AlaGlnSer 810
Db 3949 GTTCAAGCT 3957

RESULT 8
CRUPG1185 3987 bp mRNA linear ROD 27-APR-1993
LOCUS Chinese hamster p-glycoprotein mRNA (clone ADX185), complete cds.
DEFINITION
ACCESSION M59254
VERSION M59254.1 GI:191156
KEYWORDS multidrug resistance glycoprotein; p-glycoprotein.
SOURCE Chinese hamster, cDNA to mRNA, clone ADX185.
ORGANISM Cricetus sp.
REFERENCE 1. (bases 1 to 3987)
AUTHORS Devine,S.E., Hussain,A., Davide,J.P. and Melera,P.W.
TITLE Full length and alternatively spliced pgp1 transcripts in multidrug-resistant Chinese hamster lung cells
JOURNAL J. Biol. Chem. 266 (7), 4545-4555 (1991)
MEDLINE 91154265
PUBMED 1671863
FEATURES
Location/Qualifiers
1..3987
/organism="Cricetus sp."
/db_xref="taxon:10031"
/cell_type="fibroblast"
/tissue_type="lung"
1..3987
/gene="pgp-1"
/gene="pgp-1"
/codon_start=1
/product="p-glycoprotein"
/protein_id="AAA37005.1"
/db_xref="GI:191157"
translation="MTTYAVYVYTGIGAGVLIVAYIQVSWFCAAAGROIHKIRKQFFHA
IMQEGICWFDVHVGELNTRLTDDVSKINSGIDKMGFFQAMATFGGFTIGTRGW
KLTVILAIAPVLGLSAGIKWILSSFTDELQYAKAGAVAEVLAIRVIAFGGQ
KKELENNLEEARIGIKKAITANISMGAAFLIIYASALAFWYGLSVLSKEYSI
GOVLTVFVLIAPFSGQSPNIEAFANAGAYEIFNIDNKPSYDSFSKNYKPD
NIKNLEFKNIHFSPSRKDVQILKGLNLKQVQSTVALVGNSCGKSTTVQLQRLY
DPTEGVYSIDQDITRNVLREIIGVWQEPVLFTATTAEINRYGRENVTMBIEK
AVTEANAYDFIMKLPKPFEDTLVGBGAQLSGQKORIAAPALVRNPKILLDEATS
LDTESEAVQALDKAREGRTTIVIAHRLSTVRNADIIAGDGGVIVQEGNHESLMRE
KGIYFKLVMTQTAGNIELVNEVGESKNEIDNLDMSKDSASSILRRRSTRRSIRGPH

```


DDRLKSTKEALDEVPPIISFWRIKLNSSEWPFVFWGFCFCAVNGALQAPFISFISK
 VVGVRTRNDTDRKHSNLSFLILGIVSIFITFLQGGTFFRGKRGILFKRLRYMV
 FKSMRLQWDFNDPKNTGALTTLRLANDAGOVKGATGARLAVITQNTALIGTHIIS
 LLYGQOLTLILLAIVIPITAIAGVEMKLSGOALKKEKKEGSGKIATEALENPRTVV
 SLTRQKQENMYAQSLQIPYRNALKKAHVFGITFSFTQMMYFSAACFRGALVAR
 EUMTFENVLLVSALVFGMAVQVSSFAPDIYAKARVQSHIMIERKVPISIDSVTG
 GLKPNTEGNKFNVEVNYPRPDI PVLQGLNLEVKVQGTFLDLSGSGCKSTVVQL
 LERFYDPWAGTFFLDGKEVNLQVWLRHLGIVSQEPTLFCDSIAENIAYGDNRSRV
 SODEIRAANEANIHOFTESLPDKYNTRVGDGTOLSGQOKORIAIALARALVROPHILL
 LDEANSALDTESEKVVQVQALDKAREGRTCIIVIAHRLSTIQNADLLIVLVQNGKVBHGT
 HQOLLAQKGIYFSMVSVQAKR"

BASE COUNT 1194 a 822 c 944 g 1027 t
 ORIGIN Map position 1q26.

Alignment Scores:

Pred. No.: 4, 58e-176 Length: 3987
 Score: 2476.00 Matches: 468
 Percent Similarity: 77.52% Conservative: 170
 Best Local Similarity: 56.87% Mismatches: 171
 Query Match: 60.70% Indels: 14
 DB: 10 Gaps: 4

US-09-873-409-2 (1-812) x CRUPGP1185 (1-3987)

Qy 2 ValAspGluAsnAspIleArgAlaLeuAenValArgHisTyrArgAspHisIleGlyVal 21
 Db 1141 ATCGACGGCAGCAGCATCAGGACCATCAATGTGAGGTATCTGGCGGAATCATTTGGGTC 1200
 Qy 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleIleYsTyrGlyArg 41
 Db 1201 GTGAGTCAGGAACCTGTGTGTTGGCACCATCATCGCTGAAACATTCGCTATGGCCGA 1260
 Qy 42 AspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
 Db 1261 GAAATGTCTACTATGACGAGATCGGAAGCTGTCAAGGAAGCCATGCTATGACTTC 1320
 Qy 62 IleMetGluPheProAsnIlePheAsnThrLeuValGlyGluIleGlyAlaGlnMetSer 81
 Db 1321 ATCATGAAACTGCCCATAAATTTGACACTCTTTGTTGGTGAGAGAGGACACAGCTGAGT 1380
 Qy 82 GlyGlyGlnIleGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProIleIleLeu 101
 Db 1381 GGAGGACAGAAACAGAGAATCGCATTTGCTGTCGCCCTGTGTCGCCAACCCCAAGATCCTT 1440
 Qy 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerIleYsSerAlaValGlnAlaA 121
 Db 1441 TTGTTGATGAGCGGACATCAGCCTTGACACAGAAGTGAAGCCGTGTTCAGGCTGCT 1500
 Qy 122 LeuGluIleAlaSerIleGlyArgThrThrIleValAlaHisArgLeuSerThrIle 141
 Db 1501 CTGGATAAGCTAGAGAGGCGGACTACCATTTGTATAGCCACCCGCTTGTCTACAGTT 1560
 Qy 142 ArgSerAlaAspLeuIleValThrLeuIleAspGlyMetLeuAlaGluIleGlyAlaHis 161
 Db 1561 CGAAATGCTGACATCATAGCTGCGGTTTGATGTGTGTGTCATTTGTGGAGCAAGGAAATCAT 1620
 Qy 162 AlaGluLeuMetAlaIleArgGlyLeuTyrTyrSerLeuValMetSerGln----- 178
 Db 1621 GAAGAGCTCATGAGAGAGAGGCGGCAATTTCAAACTGTGTCAATGACACAGCAGCAGGA 1680
 Qy 179 -----AspIleLysAlaAspGluGlnMetGluSerMetThr 191
 Db 1681 AATGAAATTTGAATAGGAATGAAGTTGGTGAGTCTAAATAATGAAATTTGATAATTTAGAC 1740
 Qy 192 TyrSerThrGluArgIleYsThrAsnSerLeu-----ProLeuHisSerValIleSerIle 209
 Db 1741 ATGCTCTCAAAAGATTACGATCATCGAGTCTAATCAGAAGGAGATCAACTCGCAGAGATATC 1800
 Qy 210 LysSerAspPheIleAspLysAlaGluGluSerThr---GlnSerLysGluIleSerLeu 228
 Db 1801 CGTGGACCATGACCAAGCAGAGAGGCTTAGTACCAAGAGGCGCTTGGATGAAGATGTA 1860
 Qy 229 ProGluValSerLeuLeuIleLysIleLeuLysLeuAsnLysProGluTrpProPheValVal 248

Db 1861 CCTCCAAATTTCTTTTGGCGTATCTCTAAAGTTGAATTCATCTGAATGGCCATATTTTGG 1920
 Qy 249 LeuGlyThrLeuAlaSerValLeuAenGlyThrValHisProValPheSerIleIlePhe 268
 Db 1921 GTTGTATATTTCTGTGCCATGTAATAGAGCTTGCAACACGAGCATTTCTCAATAATATTT 1980
 Qy 269 AlaLysIleIleThrMetPhe---GlyAsnAsnAspLysThrThrLeuLysHisAspAla 287
 Db 1981 TCTAAGGTTGTAGGGGTTTTCACAGAATACTATGATGATGAACCAACGACATGATAGC 2040
 Qy 289 GluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMet 307
 Db 2041 AACTTGTGTTTCTTATTTCTGATCTCTGGGTCATTTCTTTTATTACATTTTCTT 2100
 Qy 308 GlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuAenHisLeuAla 327
 Db 2101 CAGGGCTTCACATTTGGCAAGCTGGAGAGATCTTCACCAACGAGCTCGGATACATGTT 2160
 Qy 328 PheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGluLysGluAsnSerThrGly 347
 Db 2161 TTCAAAATCCATGCTGAGACAGGATGTGAGCTGTTGATTAACCTTAAACCAACCACTGA 2220
 Qy 348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg 367
 Db 2221 GCATTCAGCCACCGAGCTCGCCAAACGATGCTGTCAAGTTAAAGGGGCTACAGAGCCAGA 2280
 Qy 368 IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387
 Db 2281 CTGTGCTCATTTACCCAGATATAGCAATCTTGGACAGGAATCATCATATCCCTAATC 2340
 Qy 388 TyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGly 407
 Db 2341 TATGCTGGCAGTTGACACTTCTACTCTTAGCAATGTGCCCATCATTCGCAATAGCAGA 2400
 Qy 408 MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHis 427
 Db 2401 GTGTGTTGATGAAAAATGTTGCTGGCAAGCAGCTCAAAAAGATAAGAGGAGCTAGAAGT 2460
 Qy 428 AlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArg 447
 Db 2461 TCTGGGAAGATTGCTACTGAAGCAATAGAGAACTTCGCGACTGTCTCTTTTGACTCGG 2520
 Qy 448 GluLysAlaPheGluGlnMetTyrGluMetLeuGlnThrGlnHisArgAsnThrSer 467
 Db 2521 GAGCAGAGTTTGAAATATGTATGCCAGAGCTTGCAGATACCATACAGAAATGCTCTG 2580
 Qy 468 LysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487
 Db 2581 AAGAAGACACAGCTCTTTGGGATCACAATCTCTCCACCGAGCCATGATGATATTTTCG 2640
 Qy 488 TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507
 Db 2641 TATGCTGCTGTTTTCGCGTTGCTTACTTTGGCGCAGGGAATTAATGACATTTGAA 2700
 Qy 508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeu 527
 Db 2701 AATGTTCTATTAGTATTTCTACGTATTTGTTGGTGCATGCGAGTTGGCGAGTCACT 2760
 Qy 528 ValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeu 547
 Db 2761 TCATTTGCTCTGACTATGCCAAAGCCAAAGTGTGGCATCGCATCATCATGATCAT 2820
 Qy 548 GluLysLysProAsnIleAspSerArgSerGlnGluLysLysPheProAspThrCysGlu 567
 Db 2821 GAAAGTCCCTTCCATTGACAGCTACGACCGAGGCGCTGAAGCCCTTAATACATTTGAA 2880
 Qy 568 GlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIle 587
 Db 2881 GGAATGTGAATTTAATGAAGTCGTGTTCAACTACCCACCCCGGCTGACATCCAGTG 2940
 Qy 588 LeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSer 607


```
Db 2941 CTTGAGGGCTGAACCTGGAGGTGAAGAGGGCCAGACGCTGGCCCTGGTGGGCGAGCAGT 3000
Qy 608 GlyCysGlySerThrSerValGlnLeuLeuGlnArgLeuTyAspProValGlnGly 627
Db 3001 GCGTCGGGAAGAGCAGCAGTGTGTGAGCTGTCTGAGCGGTCTTACGACCCCTGCGCGGA 3060
Qy 628 GlnValLeuPheAspGlyValAspAlaLysGluLeuLeuValGlnThrLeuArgSerGln 647
Db 3061 ACAGTGTCTTCTAGATGGCAAGAGTAACCAACTAAATGTCCAGTGGCTCGAGCGACAC 3120
Qy 648 IleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAla 667
Db 3121 CTAGGCAATGTGTCCTCCAGAGCCCATCTGTTGACTGAGCATCGCTGAGAACATCGCT 3180
Qy 668 TyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAlaAla 687
Db 3181 TATGAGACACAGCGGGTCTGTGTCAGGATGAGATCGAGAGGCGCGGAAGAGCC 3240
Qy 688 AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLys 707
Db 3241 AACATCCACCACTTCACTTGAAGTCACTGCTGACAAATACCAACACAGAGTAGGAGACAA 3300
Qy 708 GlyAlaGlnLeuSerGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuGln 727
Db 3301 GGGACTCAGCTGCGGGTGGGAGAAACAGCATCGCCATAGCAGCCCTCTGTCAGA 3360
Qy 728 LysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 747
Db 3361 CAGCTCACAATTTGCTCTGTGATGAGCAACATCAGCCCTGGATACAGAGTGAAGAG 3420
Qy 748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767
Db 3421 GTTGTCCAGGAAGCTCTGCACAAAGCCAGAGAAGCGCCACCTGCTGATGCTGCTCAC 3480
Qy 768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisGlnGlyLysIleLys 787
Db 3481 CGCGTGTCCACCATCCAGAACCGAGACTTGATCGTGTGATTCAGAAATGGCAAGGTCAAG 3540
Qy 788 GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807
Db 3541 GAGCATGGCACCACCGAGCTGTGGCAGACAAAGGCATCTATTCTCCATGCTGCTACT 3600
Qy 808 AlaGlnSer 810
Db 3601 GTGCGGCT 3609
RESULT 9
CRUPGPI 4296 bp mRNA linear ROD 22-JUN-1995
LOCUS C.griseus p-glycoprotein (isoform class I) mRNA, complete cds.
DEFINITION M60040
ACCESSION M60040
VERSION M60040.1 GI:191164
KEYWORDS multidrug resistance; p-glycoprotein; transmembrane protein.
SOURCE C.griseus adult liver and ovary, cDNA to mRNA.
ORGANISM Cricetulus griseus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
Cricetulus.
1 (bases 1 to 4296)
REFERENCE Endicott,J.A., Sarangi, F. and Ling, V.
AUTHORS Complete cDNA sequences encoding the Chinese hamster P-glycoprotein
TITLE gene family
JOURNAL DNA Seq. 2 (2), 89-101 (1991)
MEDLINE 92135886
PUBMED 1685679
FEATURES Location/Qualifiers
1..4296
/organism="Cricetulus griseus"
/db_xref="taxon:10029"
/tissue_type="liver and ovary"
/dev_stage="adult"
103..3933
/note="putative"
CDS
```

```
/codon_start=1
/product="p-glycoprotein isoform I"
/protein_id="AAA68883.1"
/db_xref="GI:191165"
/translation="MEPEDFSGRKDNFLKMGKSKKEKKEKPPVSVFTFRIYAGH
LDRLMLVGTAAIIGHVALPLMLVFGDMTDSFASVGNITPTNATNATONVNSDIFG
KLEEMTYAYVYTGIGAGVLIIVAYIOVSFWCLAAAGROIHKIROKHFHAIMNOBIFG
DVHDVGLNTRLDDVSKINEGIDKLGMEFQAMATFGGFLIGFTGKMLKLVILAI
SPVLGSAGIWKILSSFTDKELQYAKAGAAVEVLAARITVAFGGQKLEBRYNN
NLEAEKRLGKKAITANISMGAPLLIYASALAFYGTGLSVISKEYSIGQVLTVFFA
LHIFGASIGQASPNIEAFANAGAAEYIFNIIDNKPSIDFSKNGKYPDNIKLEPK
NIHFSYPSKDVOLKGLNLKVGSGQTVLVGNSCGKSTTVQLQLQYLDPTGKGVSI
DGODITINRYLREIIGVVSQEPVLPATTIAENIRYGRNVMTDELEKAVEANAYD
FIMKLPKFTLGERGAQLSGGOKRIARALVRNPKILLDLDEATSDALTESEAVV
QAALDKREGRTTIVIAHRLSTVKNADIAGFDGCVIVEQGNHHEMERKGIYFKLVN
TQADNEIELGNEVESKEINIDNMSKDSASLIRRRSTRSIRGPHDDKRLSTK
EALDEDVPIPSFWRLKLNSEWPFYVVGIFCAIVNGALQAPAFSIIFKVVGVFTNT
DDETKRHSNLSLLFLILVISPTIFLOQFTGKAGEILTKRLRYMVFKMSMLRQDV
SWFDPKNTTCAUTTRLANDAGQVKGATGARLAVITONIANLGTGIIISLILYGHQTL
LLLAIFPIIAGVVMKLSGQALKDKKELEGSGKIATEAENEFRTVSVLSTRQKPE
NMTAQSLOIPYRNALKKAVFGITFTSQAMMYFSAACFRFGAYLVARELMTFENVL
LVFSAIVFGAMAGQVSPADYAKAVKASHIIMIIEKVPISDSYTGGLKPKNTLEG
NVKNEVFNYPTRPDIPLVGLNLVKGKOTLALVSGSCGKSTTVQLQLQYLDPTDMA
GTVPLDGKEVNLNOLNOLRAHLGTVSOEPIILFDCSIAENIAYGNSVWSODETERAA
KEANIHQFISLPPKYNTRVGDGKGTQLSGGOKRIARALVRNPKILLDLDEATSDALD
TESKVVQEAALDKAREGRTCLIVIAHRLSTVKNADLIIVIQNGKVEHGHQQLLAQKG
IYFSMVSVQAGAKR"
BASE COUNT 1271 a 892 c 1029 g 1104 t
ORIGIN
Alignment Scores:
Pred. No.: 5,06e-176 Length: 4296
Score: 2476.00 Matches: 468
Percent Similarity: 77.52% Conservative: 170
Best Local Similarity: 56.87% Mismatches: 171
Query Match: 60.70% Indels: 14
DB: 10 Gaps: 4
US-09-873-409-2 (1-812) x CRUPGPI (1-4296)
Qy 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
Db 1450 ATCGACGGTCAGGACATCAGGACCATCATGTGAGTATCTCGGGAATCATTTGGGTC 1509
Qy 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
Db 1510 GTGAGTCAGGAACCTGTGTTTTCGCCACCATCTATCGCTGAAAAACATTCGCTATGCGCGA 1569
Qy 42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
Db 1570 GAAATGTCACTATGACGAGATCGAGAAAGCTGTCAAGGAAGCAACATGCTATGACTTC 1629
Qy 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
Db 1630 ATCATGAACCTCCCAATAATTGACACTCTGTGTGGAGAGAGGACACAGCTGAGT 1689
Qy 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
Db 1690 GGAGCAGACAAACAGAGATCGCCATTCGTCGCTCGCTCGCAACCCCAAGATCCTT 1749
Qy 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
Db 1750 TTGTTGGATGAGCGACATCAGCTTGGACACAGAAAGTGAAGCCGTGTTTCAGGCTGT 1809
Qy 122 LeuGluLysAlaSerLysGlyArgThrIleValValAlaHisArgGluSerThrIle 141
Db 1810 CTGGATTAAGGCTAGAGAGGCGGACTACCATTTGTATAGACCCAGCTGTGTACAGTT 1869
Qy 142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
Db 1870 CGAATGCTGACATCATAGCTGGTGTGATGTTGATGTTGATGTTGATGTTGATGTTGAT 1929
Qy 162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln----- 178
```


Db	1930	GAAGAGCTCTATGAGAGGAAGCGCATTTACTTCAAAACCTTGTCATGACACAGCAGCAGGA	1989
Qy	179	-----AspIleLysLysAlaAspGluGlnMetGluSerMetThr	191
Db	1990	AATCAAAATTGAATTAGGAATGAAGTTGGTAGTCTAAAAATGAATGTGATAATTTAGAC	2049
Qy	192	TyrSerThrGluArgLysThrAsnSerLeu-----ProLeuHisSerValLysSerIle	209
Db	2050	ATGTCCTTCAAAAGATTTCAGCATCCAGTCTAATCAGAGGAGATCAACTCGCAGAAGTATC	2109
Qy	210	LysSerAspPheIleAspLysAlaGluGluSerThr---GlnSerLysGluLysSerLeu	228
Db	2110	CGTGGACACCATCACCAGAAGAGCTTAGTACCAAGAGGCGCTTGGATGAAGATGTA	2169
Qy	229	ProGluValSerLeuLeuLysLysLeuAsnLysProGluTrpProPheValVal	248
Db	2170	CCTCAATTTCCCTTTGGCGTATCTTAAAGTTGAATTCATCTGAATGGCCATATTTGTG	2229
Qy	249	LeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePhe	268
Db	2230	GTTCGTATATCTGTGCCATAGTAATGGAGCTTGCAACACGAGCATCTCTCAATAATATTT	2289
Qy	269	AlaLysIleIleThrMetPhe---GlyAsnAsnAspLysThrThrLeuLysHisAspAla	287
Db	2290	TCTAAGCTGTAGGGGTTTTCAAGAAATATCTGATGATGAACCAACCAACGACATGATAGC	2349
Qy	288	GluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMet	307
Db	2350	AACCTGTTTTCCCTATGTTCTCGATCCCTGGGGTCATTTCTTTATTTACATTTTTTCTT	2409
Qy	308	GlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla	327
Db	2410	CAGGCTTCACATTTGGCAAGCTGGAGAGATCTCTCCACAGCGACTCCGATACATGTT	2469
Qy	328	PheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGly	347
Db	2470	TTCAAAATCCATGCTGACAGCAGGATGTGAGCTGGTTGTATAACCCCTAAAAACACCACTGGA	2529
Qy	348	GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg	367
Db	2530	GCATGTGACACACGAGCTGCGCAACGATGCTGGTCAAGTTAAAGGGGCTACAGGAGCCAGA	2589
Qy	368	IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle	387
Db	2590	CTTCCTGTCATTACCCAGATATAGCAATCTTGGGACAGGAATCATCATATCCCTAATC	2649
Qy	388	TyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGly	407
Db	2650	TATGGCTGGCAGTTGACATCTTACTCTTAGCAATGTGCCCATCATTTGCAATAGCAGGA	2709
Qy	408	MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHis	427
Db	2710	GTGTTGAGATGAATAATGTTGTCGACAAAGCAGCTTAAAGATGAAGAGGAGCTAGAAGT	2769
Qy	428	AlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArg	447
Db	2770	TCTGGAGAGATTGCTACTGAAGCAATAGAGAACTCCGCACTGCTGCTCTTTGACTCGG	2829
Qy	448	GluLysAlaPheGluGlnMetTyrGluMetLeuGlnThrGlnHisArgAsnThrSer	467
Db	2830	GAGCAGAAGTTTGAAAATATGTTATGCCAGAGCTTGCGAGATACCATACAGAAATGCTCTG	2889
Qy	468	LysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla	487
Db	2890	AAGAAAGACACGCTGTTGGGATCACATTTCTCCTTCACCCAGGCGCATGATGATTTTTTCG	2949
Qy	488	TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu	507
Db	2950	TATGCTCTTGTTCCTGGTTTGGTGCTTACTTGGTGGCAGGGGAATTAATGACATTTGAA	3009
Qy	508	GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeu	527

Db	3010	AAATGCTTATTAGTATCTCAGCTATTGCTTTGGTGCCATGGCAGTTGGCAGGTCAGT	3065
Qy	528	ValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeu	547
Db	3070	TCATTTCCTCTGACTATGCCAAGCCAAAGTGTGGCATCGCATATCATCATGATCATT	3129
Qy	548	GluLysLysProAsnLeuAspSerArgSerGlnGluGlyLysLysProAspThrCysGlu	567
Db	3130	GAAAGAAGTCCCTTCCATTGACAGCTACAGCACCGGAGGCGCTGAAGCCTTAATACATTGGA	3189
Qy	568	GlyAsnLeuGluPheArgGluValSerPheThrProCysArgProAspValPheIle	587
Db	3190	GGAAATGTGAATTTAATGAAGTCGTGTTCAACTACCCACCAGCCCTCGACATCCAGCT	3249
Qy	588	LeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSer	607
Db	3250	CTTCAGGGCTGAACCTGGAGGTGAAGAGGCCAGACGCTGGCCCTGTTGGCAGCAGCT	3309
Qy	608	GlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGly	627
Db	3310	GGCTTCGGGAAGAGACAGTGGTTCAGCTGCTCGAGCGGTCTTACAGCCCATGCCCGGA	3369
Qy	628	GlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGln	647
Db	3370	ACAGTGTTCCTAGATGGCAAGAAAGTAACCAACCTAAATGTCAGTGGCTCCGAGCACAC	3429
Qy	648	IleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAla	667
Db	3430	CTAGCGATTGTCTCCAGAGCCCATCTCTGTTTGTACTGCAGCATCGCTGAGAACATCGCT	3489
Qy	668	TyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAla	687
Db	3490	TATGGACACAACAGCCGGTCTGTGCCAGGATGAGTCGAGAGGGCGCGCAAGGAAGCC	3549
Qy	688	AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLys	707
Db	3550	AACATCCACCAAGTTCATTGAGTCACTGCCCTGCAAAATACACACAGAGTAGGACACAA	3609
Qy	708	GlyAlaGlnLeuSerGlyGlyLysGlnArgLeuAlaIleAlaArgAlaLeuGln	727
Db	3610	GGGACTCAGCTGTCTGGTGGGCGAAGCAACCGCATCGCATAGCAGCGCCCTCGTCAGA	3668
Qy	728	LysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys	747
Db	3670	CAGCCTCACATTTTGTCTCTGGATGAAGCAACATCAGCCCTGGATACAGAGAGTCAAAAG	3729
Qy	748	ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis	767
Db	3730	GTTCTGCAGAAAGCTCTGGACAAGCCAGAAAGCCCGCCCTGCTATTTGATCGCTCCAC	3789
Qy	768	ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys	787
Db	3790	CGCTGTCCACCATCCGAACCGAGACTTGATCTGGTGATTCAGAAATGGCAGAGTCAAG	3849
Qy	788	GluGlnGlyThrHisGlnGlnLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn	807
Db	3850	GAGCATGGCACCACAGCAGCTGCTGGCCAGACAGAAAGGCATCTATTTCTCCATGGTCAGT	3909
Qy	808	AlaGlnSer 810	
Db	3910	GTGCAGGCT 3918	
RESULT 10			
CRUPG1165			
LOCUS			
DEFINITION Chinese hamster p-glycoprotein mRNA (clone ADX165), complete cds			
ACCESSION M59253			
VERSION M59253.1 GI:191154			
KEYWORDS multidrug resistance glycoprotein; p-glycoprotein.			
SOURCE Chinese hamster, CDNA to mRNA, clone ADX165.			
ORGANISM Cricetus sp.			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae			
CRUPG1165			
LINEAR ROD 27-APR-			

VERSION M59253.1 GI:191154
KEYWORDS multidrug resistance glycoprotein; p-glycoprotein.
SOURCE Chinese hamster, cDNA to mRNA, clone AUX165.
ORGANISM *Cricetulus* sp.

REFERENCE 1 (bases 1 to 4304)
 AUTHORS Devine,S.E., Hussein,A., Davide,J.P. and Melera,P.W.
 TITLE Full length and alternatively spliced pgpi transcripts in multidrug-resistant Chinese hamster lung cells
 J. Biol. Chem. 266 (7), 4545-4555 (1991)
 MEDLINE 91154285
 PUBMED 1671863

FEATURES
 source
 1..4304
 /organism="Cricetulus sp."
 /db_xref="taxon:10031"
 /cell_type="fibroblast"
 /tissue_type="lung"
 1..4304
 /gene="pgp-1"
 111..3941
 /gene="pgp-1"
 /codon_start=1
 /product="p-glycoprotein"
 /protein_id="AAA37004.1"
 /db_xref="GI:191155"
 /translation="MEPEDEFSGRKDKNFKMGRKSKKKEKKPPVSVFTMFRVAGWLDRLMYLGTAAIHGHVALPLMLVFGDMTDSFASVGNIPNTATNNAQVNASDIFGKLEEMTYAYTIGAGLVIAIYQSFVCLAGRQIHKIRQKPFHAINWOBIGWFDVHVGELNRLTDDVSKINEGIDKIGMFFQAMATPFGFIIGFTRGWKLTLVLAISPVGLSAGIAWAKILSSFTDKELQAYAKAGAAVEVLAARIVIAFGQKKLELYNNNLEAEKRGIKKAITANTSMGAAPLLIYASVALAFWYGTSLVISEYIGOVLTVPFAVLIAFPGISQASPNTEAFANARGAAVEIFNIDNKPSIDSFKNQYKPDNIKGNLEPKNIHESYPRKDVQILKGLNLKQSGQTVLVNSCGSKSTTVQLLQRLYDTEGVVSI DQDILRTNRYRLIEIGVSVQEPVLFATTAENIRYKRENTVMDIEBAKVENAYDPIMKLPHKFDLTVGRGAQGGQKQRIATARALVNPXKILLDETSALDPTSEAVYQAALDKAREGRTIVIAHRLSTVRNADIAGFDGGVIEQGNHLELMREKGIYFKLVMTOTAGNETELGNEVESKNEIDNLSKSDASSLIRRRSTRSRIRGPHDODRKLSTKEEALDEDVPIISFWRILKLSSEWPFVVGIEICAIWNGALOPAFSIIISFKVGVFTRNTDDETRKDSNFSLLFLILGVISFIITFLOQFTFGKAGEIITKRLRYVWFKSMLRQDVSWDFPNKNTGALTIRLANDAGQVKGTGARLAVITQNIANLGTIIISLIYQWLTLLLAALVPIAIGAVVEMKLSQALKDKKELEGSKIAATEAENFRTVSVLTREQKFNMYAQSLQIPYRNALKKHAHFGITPFTQAMMYFSYAACFRGAIVLARELMTFENVLVFSAIVFGMAVGVSFPADYAKAKVSASHIMIIEKVPISIDSYSGTGLKPNLTLEGNVKNEVVPNTPRDIPVLOGLNLEVKGGOTLALVSGCGCKSTVQVLLERFDPMA GTVFJUDGKQVNLQWMLRHIGIVSQPIFLDCGSAENIAYGNSRVVSQDEIERAA KEANTHQIESLPDKYNTRVGDKGTQLSGGQKRIATARALVROPHILLDEATSLD TESERKVOEALDKAREGRTCVIAHRLSTQVNIADLIIVVIONGKHEHGHQOLLAQKG IYFSMVSVOAGAKR"

BASE COUNT 1275 a 895 c 1031 g 1103 t
 ORIGIN Map position 1q26.

Alignment Scores:
 Pred. No.: 5,07e-176 Length: 4304
 Score: 2476.00 Matches: 468
 Percent Similarity: 77.52% Conservative: 170
 Best Local Similarity: 56.87% Mismatches: 171
 Query Match: 60.70% Indels: 14
 DB: 10 Gaps: 4

US-09-873-409-2 (1-812) x CRUPGP1165 (1-4304)

Qy 2 ValAspGluAsnAspIleArgAlaLeuAenValArgHisTyrArgAspHisIleGlyVal 21
 Db 1458 ATCGACGGACAGACATCAGGACCAATCAATGTGAGTATCTCGCGGAAATCAATCGGGGTC 1517

Qy 22 ValSerGlnGluProValLeuPheGlyThrIleSerAsnAsnIleLysTyrGlyAsp 41
 Db 1518 GTGAGTCAGAACCTGTGTTTGGCACACATACCTATCGCTGAAACAACTTCGCTATGGCCGA 1577

Qy 42 AspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
 Db 1578 GAAATGTCTACTATGACGAGATCGAGAAAGCTGTCAAGGAAGCAATGCCATGACTTC 1637

Qy 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
 Db 1638 ATCATGAAACTGCCCCATAAATTGACACTCTTTGTTGGTGAGAGAGGACACAGCTGAGT 1697

Qy 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
 Db 1698 GGAGACAGAAACACAGAAATGCCATTGCTGTCGCTCGCCCAACCCCAAGATCCTT 1757

Qy 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
 Db 1758 TTGTTGGATGAGGCGACATCAGCCTTGACAGAAAGTGAAGCCGTGGTTCAGGCTGT 1817

Qy 122 LeuGluLysAlaSerLysGlyArgThrThrIleValAlaAlaHisArgLeuSerThrIle 141
 Db 1818 CTGGATAGGCTAGAGAGGCGGACTACCATTTGTGATAGCCACCGCTTGTCTACAGTT 1877

Qy 142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
 Db 1878 CGAAATGCTGACATCATGCTGGTGTGATGCTGTGTCATTGTGAGCAGCAAGAAATCAT 1937

Qy 162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerCln----- 178
 Db 1938 GAAGAGCTCATGAGAGAGAGGCGCATTTTCAAACTTGTCTATGACACAGACAGCAGGA 1997

Qy 179 -----AspIleLysLysAlaAspGluGlnMetGluSerMetThr 191
 Db 1998 AATGAATTTGAATAGGAATGAAGTTCGTGAGTCTAAATAATGAATTCATAATTAGAC 2057

Qy 192 TyrSerThrGluArgLysThrAsnSerLeu-----ProLeuHisSerValLysSerIle 209
 Db 2058 ATGTCTCTCAAAAGATTTCAGCATCCAGTCTAATCAGAAGGAGATCAATCGCAGAAAGTATC 2117

Qy 210 LysSerAppPheIleAspLysAlaGluGluSerThr--GlnSerLysGluLysSerLeu 228
 Db 2118 CGTGACACCATGACACAGAGGAGGCTTAGTCAAAAGAGCGCTTGGATGAAGATGTA 2177

Qy 229 ProGluValSerLeuLysIleLeuLysLysLeuAsnLysProGluTyrProPheValVal 248
 Db 2178 CTTCAATTTCTTTTGGCGTATCTTAAGTTGAATTCATCTGAATGGCCATATTTTGG 2237

Qy 249 LeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePhe 268
 Db 2238 GTTGGTATATCTGTGCCATAGTAATAAGGAGCTTGCAACACGACATCTCTCAATAATAT 2297

Qy 269 AlaLysIleIleThrMetPhe--GlyAsnAsnAspLysThrThrLeuLysHisAspAla 287
 Db 2298 TCTAAGTTGTAGGGGTTTTCAAGAAATACTGATGATGAACCAACCAACGACATGATAC 2357

Qy 288 GluIleThrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMet 307
 Db 2358 AACTTGTCTTCTTATGTTCTGATCTTGGGGTCATTTCTTTTATCATTTTTTCTT 2417

Qy 308 GlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla 327
 Db 2418 CAGGCTTCACATTTTGGCAAAAGCTGGAGAGATCTCTCACCAGCGACTCCGATACATGTT 2477

Qy 328 PheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGluLysGluAsnSerThrGly 347
 Db 2478 TTCAAAATCCATGCTGAGACAGATGTCGTTGATTAACCCCTTAAACCAACCACTGGA 2537

Qy 348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyValaThrGlySerArg 367
 Db 2538 GCATTGACACACAGGCTGCCAACCATGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 2597

Qy 368 IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387
 Db 2598 CTTGCTGTCATTACCCAGAAATATACCAATCTTGGGACAGGAATCATCATATCCCTAATC 2657

Qy 388 TyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGly 407
 Db 2658 TATGCTGTCAGTTGACACTTCTACTCTTAGCAATTTGTCCTCCCATCATTCGTAATGAG 2717

Qy 408 MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLysHis 427
 Db 2718 GTGGTTGATGATAAAATGTTGCTGGCAAGCACTTAAAGATGAAGAGGAGCTAGAGGT 2777


```
QY 428 AlaGlyIysIleAlaThrGluAlaLeuGluAAsnIleArgThrIleValSerLeuThrArg 447
Db 2778 TCTGGGAAGATTCTACTGAAGCAATAGAGAACTCCGACGTCTCTCTTTGACATCG 2837
QY 448 GluLeuAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSer 467
Db 2838 GAGCAGAAGTTTGAAATATGTATGTCAGAGCTTGCAGATACCACATACAGAAATGCTCTG 2897
QY 468 LysLeuAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487
Db 2898 AAGAAAGCACACGCTTTGGATACATCTCTCTCCAGCCAGCCATGATGATTTTTCG 2957
QY 488 TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507
Db 2958 TATGCTGCTTTGTTTGGGTTTGGTCTTACTTGGTGGCAGCGGAATTAATGACATTGAA 3017
QY 508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeu 527
Db 3018 AATGTTCTATTAGTATCTCAGCTATGTTCTTGGTGCCATGGCAGTGGGCGAGTCACT 3077
QY 528 ValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeu 547
Db 3078 TCATTTGCTCTCTGACATGCAAGCAAGTCTCGCATCGCATATCATCATGATCAT 3137
QY 548 GluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGlu 567
Db 3138 GAAAGAGTCCCTTCCATTCAGACGTACAGCACCGGAGCGCTGAAGCCTAATACATTGGA 3197
QY 568 GlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIle 587
Db 3198 GGAATATGAAATTTAATGAAGTCGTGTTCACTCCACCCGCGCTGACATCCAGTG 3257
QY 588 LeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSer 607
Db 3258 CTTTCAGGGCTGAACCTGGAGGTGAAGAGGCGCAGACGCTGCGCTGGGCGCACAGT 3317
QY 608 GlyCysGlyLysSerThrSerValGlnLeuGlnLeuGlnArgLeuTyrAspProValGlnGly 627
Db 3318 GGCTGGGGGAAGACACAGTGGTCCAGCTCTCGAGCGGTCTTACGACCCCATGGCGCGA 3377
QY 628 GlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnThrLeuArgSerGln 647
Db 3378 ACAGTGTCTTAGTGGCAAGAAAGTAACCAACTAATGTCAGTGGCTCCGAGGACAC 3437
QY 648 IleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAla 667
Db 3438 CTAGGACATTGTGTCACAGAGCCCTCTCTGTTGACTGCGACATCCTCGAGAATCGCT 3497
QY 668 TyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAlaAla 687
Db 3498 TATGGAGACAAACAGCGGGTGTGTCCAGGATGAGATCCAGAGGGCGGCGAAGAGCC 3557
QY 688 AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLys 707
Db 3558 AACATCCACAGTTCATTAGTCACTGCTCGCAAAATACAACACAGGATGAGAGACAA 3617
QY 708 GlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuGln 727
Db 3618 GGGACTCAGCTGTGCGGTGGCGAAGAACACGATCGCCATAGCAGCGCCCTCGTCAGA 3677
QY 728 LysProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 747
Db 3678 CAGCCTCACATTTGCTCTCGGATGAAGCAACATCAGCCCTCGATATGAGTGAAGAAAG 3737
QY 748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767
Db 3738 GTTGTCCAGGAAGCTTGGCAAGGCGAGAGGCGCACCTCGATGATGATCGCTCAC 3797
QY 768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys 787
Db 3798 CGCCTGTCACCATCCAGAACGAGACTTGTGTTGATGATGATGATGATGATGATGATGAT 3857
QY 788 GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807
```

```
Db 3858 GAGCATGGCACCACGACGCTGCTGGCACAGAAAGCATCTATTTCCTCCAGGTCACT 3917
QY 808 AlaGlnSer 810
Db 3918 GTGCAGGCT 3926
RESULT 11
OAU78609 3858 bp mRNA linear MAM 11-OCT-2001
LOCUS Ovis aries multidrug resistance protein-1 (mdr1) mRNA, complete
DEFINITION
ACCESSION U78609
VERSION U78609.1 GI:2149086
KEYWORDS
SOURCE Ovis aries.
ORGANISM Ovis aries
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Caprinae; Ovis.
REFERENCE
1 (bases 1 to 3858)
AUTHORS Longley, M., Phua, S.H., van Stijn, T.C. and Crawford, A.M.
TITLE Isolation and mapping of the first ruminant multidrug resistance
genes
JOURNAL Anim. Genet. 30 (3), 207-210 (1999)
MEDLINE 99371931
PUBMED 10442984
REFERENCE
2 (bases 1 to 3858)
AUTHORS Longley, M. and Crawford, A.M.
TITLE Direct Submission
JOURNAL Submitted (18-NOV-1996) Biochemistry Department, University of
Otago, Dunedin, New Zealand
FEATURES
Location/Qualifiers
1..3858
/organism="Ovis aries"
/db_xref="taxon:9940"
gene 1..3858
CDS 1..3858
/gene="mdr1"
/codon_start=1
/product="multidrug resistance protein-1"
/protein_id="AAB58489.1"
/db_xref="GI:2149087"
/translation="MDLEGDRNGRAAGNFLKRDKRPFSSKDEKKEKRPVSTPTMF
RYSNWLRLCMVLGTAAI IHGAGLPLMLVFGDMTDSFAGAGLNGNTLNSINTST
IDREYCKLEKEMTVAYYSGAGVLI AAYLOVSPWCLAAQROVHRIKQFPHAI
MQSIGPDVHDVCKLNLRLNDVSKINEGIDKIMFQMAVTFGTFGTTGN
LTLVILAIPLVLSAAIWAIISSFTDDKLLAYAKAGAAVEVLAALKTIVAFGGQK
KELERYKNLEAKRIGIKKAITANI SMGAFLITLYASALFYAGTSLYSREYSIG
QVLTVFSLVIGAFSIGQSPNIEAFANARGAAVEFKI IDNKPSIDSYNTGHPDN
IKGNLEPRNVHFYPSRNEVKILKGLNLKVGSGOTVALVNGSGCKSTTVOLMORLYD
PTEGMVSDIGDITINRVYREIIGVVSORPVLFAITAEINRYRGREDVTMDEIOKA
VKEANADYFIMKLNFKFDLTVGERQAQSGQKQRIARALVRNPKILLDEATSAI
GTESEVQAALDKARKGRITTVIAHRLSTVNRNADIAGLDGVIVEEGSHDLMGKR
GIYKFLVTMQTKNELESTPESLSTDDLYMSQDSRSLRRKTRRSIRGSQS
QDKLSTEETLDESVPVPSFWRLKLNITETWPFVGVFCAI INGALOPAFSVFSRI
IGITRNNDETKQNSLRLFLIGIISFITFLQGTFFGKAGELTTRRLRYLVE
RSMLRQVSWFDDPKNTTGALTTLANDAAOVKAGVSR LAVITQNIANLTGIIISL
IYGHQUTLLILAIPIIAVAGVIEKMLSGQALKKKELEGAKIATEIENFRIVS
LTRREERFEYMQALQVYPYRNSLRRAHVFGITFTIQAMYFVAGCFRFGAILVAQG
IMEFQDVLIVFSAVVFGMAVQVSTPADYAKAKVSAHVINIIEKTLPLDSVSTEG
LKPTVSGSVAFNDVFNVPYRDPVPLRHLGSLVKVKGQTLALVSGSCGSKSTVQLL
ERFYDPLAGTVFIDGKEVKQLNVOLRAHMGIVSQEPLFDPCSTIGENAYGNSRVVS
OEEIHEAKENIHSFTLEMLPKYNTVRGDKTOLSGOKORIAARALVRQPHILL
DEATSAIDTESEKVVQEAIDKAREGRTCTIVAHRLSTIQNALDILVVFQNGRIKEGTH
COLLAQKGIYFTVWSVQGRKQ"
```

BASE COUNT 1072 a 853 c 1002 g 931 t

Alignment Scores: 1.74e-175 Length: 3858
Pred. No.: 2468.00 Matches: 475
Score:

Percent Similarity: 76.73% Conservative: 158
Best Local Similarity: 57.58% Mismatches: 174
Query Match: 60.51% Indels: 18
DB: 4 Gaps: 4

US-09-873-409-2 (1-812) x OAU78609 (1-3858)

Qy 2 ValAspGluAsnAspIleArgAlaLeuAenValArgHisTyArgAspHisIleGlyVal 21
Db 1372 ATCGACGGACAGGATATCAGGACCACTCAACGTAAGGATCTCGGGGATATTTCGGGGTG 1431
Qy 22 ValSerGlnGluProValLeuPheGlyThrIleSerAsnAsnIleIysTyArg 41
Db 1432 GTGAGTCAGAGCCCTGTGCTGTTGCCACACGATAGCTGGAACAATTCGCTACGGCCGT 1491
Qy 42 AspAspValThrAspGluMetGluArgAlaAArgGluAlaAsnAlaValThrAspPhe 61
Db 1492 GAAGATGTCACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1551
Qy 62 IleMetGluPheProAsnIysPheAsnThrLeuValGlyGlyGlyAlaGlnMetSer 81
Db 1552 ATCATGAACTGCCTAATAAGTTTGACACCTGTTGGAGAGAGAGAGAGAGAGAGAGAG 1611
Qy 82 GlyGlyGlnIysGlnArgIleAlaAArgAlaLeuValArgAsnProIysIleLeu 101
Db 1612 GTGACAGACAGAAAGAAATGCCATTTGCCCGGCCCTGGTTGCCAACCCTCAAGATCCTC 1671
Qy 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaA 121
Db 1672 CTGTTGGATGAGCCACCTCAGCGTGGCCACCGAGAGTGAAGCGGTGTTACGGCGCC 1731
Qy 122 LeuGluLysAlaSerIysGlyArgThrIleValValAlaHisArgLeuSerThrIle 141
Db 1732 CTGGAATAGCCAGAAAGCCGACCACTTGTGATAGCGCATCGTGTCCACGGTT 1791
Qy 142 ArgSerAlaAspIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
Db 1792 CCCAATGCTGCATCATCTGCTGTGATGATGATGATGATGATGATGATGATGATGAT 1851
Qy 162 AlaGluLeuMetAlaLysArgGlyLeuTyThrSerLeuValMetSerGlnAspIleLys 181
Db 1852 GATGAGCTCATGGGAAAGAGCATTTACTTCAAACTTGTACAACTGTCACATGCG--ACAAA 1908
Qy 182 LysAlaAspGluGlnMetGluSerMetThrTyThrSerThrGluArgLysThrAsnSerLeu 201
Db 1909 CGAAATGAACCTGAAATGGAAATATCCCTGTTGATGATGATGATGATGATGATGATG 1968
Qy 202 ProLeuHisSerValLysSerIleLysSerAspPheIle----- 214
Db 1969 TACATGCTCCCAAGATTCA---AGATCCAGTCTAATCAGAAAGAAATCAACTCGCAGG 2025
Qy 215 -----ApLysAlaGluSerThrGlnSerLysGluIle 226
Db 2026 AGTATCGTGGATCCCAAGCCAGGACAGAAAGCTCAGTACAGAAAGAACTTTGGATGAA 2085
Qy 227 SerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPhe 246
Db 2086 AGTGATCTCCCGTTCTTTTCGAGGATCTGAAGCTGAATATAACTGAATGGCCTTAT 2145
Qy 247 ValValLeuGlyThrLeuAlaSerValLeuAenGlyThrValHisProValPheSerIle 266
Db 2146 TTTTGGTTGGTGATTTTGGCGCCATATATAAATGGAGCTCTGCAACCAGCAATTCCTCAGTC 2205
Qy 267 IlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLys---ThrThrLeuLysHis 285
Db 2206 ATATTTCAAGGATATAGGAATCTTCCACAGAAATGACAAATGACAGAAACCAACAGCAG 2265
Qy 286 AspAlaGluIleTySerMetIlePheValIleLeuGlyValIleCysPheValSerTy 305
Db 2266 AACAGTAACCTGTTTCACTGTTGTTCTTAATCCTTGGAAATATTCTTTTATTACATTT 2325
Qy 306 PheMetGlnGlyLeuPheTyArgAlaGlyGluIleLeuThrMetArgLeuArgHis 325
Db 3406 ATCGCTACGGAGACACAGCCGGTCTGTCTCCAGGAGGAGATTGAGCACCGCCCAAG 3465

Db 2326 TTCCTTACGGGCTTTCACGTTTCGGGAAACAGCGCGAGATCTCTCACAGGGGGTTCGGGTAC 2385
Qy 326 LeuAlaPheLysAlaMetLeuTyArgGlnAspIleAlaTrpPheAspGluLysGlnAsnSer 345
Db 2386 CTGGTTTTCAGTCGATCTGAGACAGGATGTGAGCTGGTTCGATGACCCATAAACACC 2445
Qy 346 ThrGlyGlyLeuThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGly 365
Db 2446 ACGGAGCATTTGACACACGAGCTTCCCAATGACGCGCTCAAGTTTAAAGGGGCTTAGGT 2505
Qy 366 SerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSer 385
Db 2506 TCCAGCTTGTGTGTTATCCAGCAATATTGCAAACTTTGGCAAGGCAATTATCATATCC 2565
Qy 386 PheIleTyArgTyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaVal 405
Db 2566 CTCATCTATGCTGGCAGTTAACACTTCTCCTGTTAGCAATTTGACCATCATTCAGTAT 2625
Qy 406 ThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeu 425
Db 2626 GCAGGAGTTATTGAAATGAAATGCTGTGGAACAAGCCCTGAAAGATAAGAAAGAACTG 2685
Qy 426 LysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeu 445
Db 2686 GAAGTGTCTGGAAATGCTGCTGGAACAAGCCCTGAAAGATAAGAAAGAACTG 2745
Qy 446 ThrArgGluLysAlaPheGluGlnMetTyArgGluGluMetLeuGlnThrGlnHisArgAsn 465
Db 2746 ACTCGGAGGAGAGGTTTGAATATATATATATATATATATATATATATATATATATAT 2805
Qy 466 ThrSerLysLysAlaGlnIleGlySerCysTyArgAlaPheSerHisAlaPheIleTy 485
Db 2806 TCCTTGAGGAGAGCGCATGCTTTCGAAATATACATTTCCATCCTCAGGCAATGATGAT 2865
Qy 486 PheAlaTyArgAlaAlaGlyPheArgPheGlyAlaTyLeuIleGlnAlaGlyArgMetThr 505
Db 2866 TTCTCTATGCTGGCTGTTCCGGTTTGGTCTACTTGTGGCCAGGCAATCATGGAG 2925
Qy 506 ProGluGlyMetPheIleValPheThrAlaIleAlaTyArgGlyAlaMetAlaIleGlyLys 525
Db 2926 TTTTCAGGATGTTCTCTAGTGTCTCAGCTGTGTCTTGTGGTCCCATGGCTGGGACAG 2985
Qy 526 ThrLeuValLeuAlaProGluTySerLysAlaLysSerGlyAlaAlaHisLeuPheAla 545
Db 2986 GTCACTACATTTGCTCTGACTATGCCAAGGCAAAAGTGTCAAGCAGCCACGCTCATCA 3045
Qy 546 LeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThr 565
Db 3046 ATCATTTGAAAAAATTCCTCTGATTGACAGCTACAGCAGGAAAGGCTTAAAGCCGAGTACA 3105
Qy 566 CysGluGlyAsnLeuGluPheArgGluValSerPheTyProCysArgProAspVal 585
Db 3106 GTGGAAAGAGTGTGGCATTTAATGACGCTGTGTTCAACTACCCCATCCGACAGACGTC 3165
Qy 586 PheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGly 605
Db 3166 CCAGTGCTTCGGGGCTGAGCCTGAGGCTGGAAGAGGCGCAGACGCTGGCGCTGGTGGCG 3225
Qy 606 SerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyArgProVal 625
Db 3226 AGTAGTGTGTGGGAAGAGCAGCGTCTGCTCCAGCTCTCCAGGCGGTTCTATGACCCCTTG 3285
Qy 626 GlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArg 645
Db 3286 GCTGGCAGAGTGTATTGACGGCAAGAGAGTAAGAGCAGCTGAATGTCCAGTGGCTGGCG 3345
Qy 646 SerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsn 665
Db 3346 GCACACATGGGCATCGTCTCCAGGAGCCCATCTCTGTTGACTGCAGCATCGCGAGAAT 3405
Qy 666 IleAlaTyArgLysAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsn 685
Db 3406 ATCGCTACGGAGACACAGCCGGTCTGTCTCCAGGAGGAGATTGAGCACCGCCCAAG 3465

Qy	162	AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln-----	178
Db	2027	AGTGAGCTGATGAAGAAGGAGTCTACTTACAGCTCGTTAACTACGACAGCAGCA	2086
Qy	179	-----AspIleuValLysAlaAspGluGlnMetGluSerMet	190
Db	2087	AGCCAGATCCTGTCAGAAGAAATTTGAAGTTGAGCTAAGTGACGAAGAAGGCTGCTGGAGAT	2146
Qy	191	ThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLys	210
Db	2147	GTGSCCCCAAAATGGCTGGAAGACGACGATATTTAGGAATTTCTACAAAGAAAAGTCTTTAA	2206
Qy	211	SerAspPheIleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGlu	230
Db	2207	AGTCCACATCAGAAATAGCTGGATGAAGAAACCAATAGAACTTGATGCAAACGTCGCCACCA	2266
Qy	231	ValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTyrProPheValValLeuGly	250
Db	2267	GTGTCTTTCTGAAGGTCTTAAACTGAAATAAACAAGAGTGGCCCTACTTGTGTGGGGGA	2326
Qy	251	ThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLys	270
Db	2327	ACAGTCGTGCCATTGCCAATGAGCGCCTCCAGCGCGCTTCTCCATCATCTGCTCTGAG	2386
Qy	271	IleIleThrMetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyr	290
Db	2387	ATGATAGCTATCTTTGGCCCTGGGGATGACGACGACGAACAGCAGCAAAAGTGTAACTGTC	2446
Qy	291	SerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeu	310
Db	2447	TCCCTGTGCTCTTTGGGCCCTAGAGAGTCTCTCCTCTTACTTCTTCTTCCCTTACGGGCTTC	2506
Qy	311	PheTyrGlyArgAlaGlyLeuLeuThrMetArgLeuArgHisLeuAlaPheLysAla	330
Db	2507	ACGTTTGGGAAAGCTGGAGAGATCCTCACCAACAGGCTCCGGTCCATGCGCCTTTAAAGCG	2566
Qy	331	MetLeuTyrGlnAspIleAlaTyrPheAspGluLysGluAsnSerThrGlyGlyLeuThr	350
Db	2567	ATGCTAAGGCAGGACATAGCTGGTTGATGATGATATAAAACAGTACTGGAGCACTTTCT	2626
Qy	351	ThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyVal	370
Db	2627	ACAAGACTCGCCACAGATCTGCGCAAGTCCAAAGAGAGCCACGGGAACCAAGTTGGCTTTTA	2686
Qy	371	LeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyr	390
Db	2687	ATTGCACAGAACACAGCCACCTTGGAACCGGTATTATTATATCATTTATTACGGTGG	2746
Qy	391	GluMetThrPheIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGlu	410
Db	2747	CAACTGACACTTCTGCTGTTATCGGTTGTTTCCATTCAATGCTGTGACAGGAATTTGTTCAA	2806
Qy	411	ThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLys	430
Db	2807	ATGAAATGTTGGCTGCCAATGCCAGAGAGATATAAAGAAATCGAAGCTGCTGGAAAG	2866
Qy	431	IleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGlyLysAla	450
Db	2867	ATTGCAACAGAGGCATAGAAAATATTGNACTGTGTATCTCTGACCCAGAAAGAAA	2926
Qy	451	PheGluGlnMetTyrGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAla	470
Db	2927	TTTGAGTCAATGATGTTGAAAATTTGCATGGACCTTACAGGAAATTCGGTGGGAAGCA	2986
Qy	471	GlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAla	490
Db	2987	CACATCTACGCGCATCACTTTTAGCATCTCCCAAGCATTCATGTATTTTTTCTATTGCTGGC	3046
Qy	491	GlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPhe	510
Db	3047	TGTTTTGATTTGGTCTTTCTACCTAAATTTGGAATGGACATATCGCTTCAAGATGTCATT	3106
Qy	511	IleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAla	530

[illegible]


```

FEATURES      STREMMEL WOLFGANG (DE)
source        Location/Qualifiers
1..3924      /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /note="MDR3/menschliche MDR-Genfamilie"
BASE COUNT   1145 a 790 c 977 g 1012 t
ORIGIN

Alignment Scores:
Pred. No.:    2,74e-175      Length:      3924
Score:         2465.50       Matches:     478
Percent Similarity: 76.46%   Conservative: 149
Best Local Similarity: 58.29% Mismatches:    182
Query Match:    60.44%      Indels:       11
DB:             6           Gaps:         3

US-09-873-409-2 (1-812) x AX024455 (1-3924)
QY 2 ValAspGluAsnAspIleArgAlaLeuAenValArgHisTyrArgAspHisIleGlyVal 21
DB 1395 ATTGATGGGCAGGATATTAGGAATTTAAATGTAACCTATCTGAGCGGAATCATTTGGTGTG 1454
QY 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
DB 1455 GTGAGTCAGAGCGCGTGTGTTTCCACCACAAATTCCTGGAATAATTTGTTATGGCCGT 1514
QY 42 AspAspValThrAspGluMetGluAtrGAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
DB 1515 GGAAATGTAAACCATGATGATGAATAAGAAAGCTGTCAAGAGGCCAACCGCTATGAGTTT 1574
QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
DB 1575 ATCATGAATTTACACAGAAATTTGACACCCCTGTTGGAGAGAGGGGCCAGCTGTAGT 1634
QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaAtrGAlaLeuValArgAsnProLysIleLeu 101
DB 1635 GGTGGGAGAGAGGAGGATCGCCATTTGCAGTGCCTGTTGCAACCCCAAGATCCCTT 1694
QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
DB 1695 CTCCTGATGAGCCACGTCAGCATTTGGACACAGAAAGTGAAGCTGAGGTACAGGCAGCT 1754
QY 122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgIleSerThrIle 141
DB 1755 CTGGATAAGCCACAGAGAGCCGACCATTTGTGATAGCACACCGCACTGTCTACGGTC 1814
QY 142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
DB 1815 CGAAATGCAGATGTCATCGCTGGGTTTGAGGTGGAGTAAATTTGTGGAGCAAGAGCCAC 1874
QY 162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuVal----- 175
DB 1875 AGCGAACTGATGAAGAGGAGGAGGGGTGACTTCAAACTGTGCAACATGCAGACATCAGGA 1934
QY 176 -----MetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyr 192
DB 1935 AGCCAGATCCAGTCAGAGAATTTGAACATAATGATGAAGGAGCTGCCACTAGAAATGGCC 1994
QY 193 SerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLys----Ser 211
DB 1995 CCAAAATGGCTGGAAATCTCGCCCTATTAGGCATTTCTACTCAGAAACACCTTAAAAATTCA 2054
QY 212 AspPheIleAspLysAla---GluGluSerThrGlnSerLysGluIleSerLeuProGlu 230
DB 2055 CAAATGTGTCAGAAAGAGCCTTGATGTGGAAACCGATGGACTGGAACAAATGTCACCA 2114
QY 231 ValSerLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGly 250
DB 2115 GTGTCCTTTCTGAAGGTCCTGAAACTGAAATGAAATGAAACAGAAATGGCCCTACTTTGTCGGGA 2174
QY 251 ThrLeuAlaSerValLeuAanGlyThrValHisProValPheSerIleIlePheAlaLys 270

```

```

DB 2175 ACAGTATGTGCCATTGGCCAAATGGGGGGCTTCAGCCGGCATTTTTCAGTCATATCTCAGAG 2234
QY 271 IleIleThrMetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyr 290
DB 2235 ATCATAGCGATTTTGGACCGCGCATGATGCAAGAGGAGCAGACAGAGTGCACATATTC 2294
QY 291 SerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeu 310
DB 2295 TCTTTGATTTTCTTATTTCTGGGAATTTCTTTTCTTTTCTTTCTTTCTTTCTTTCTTTCT 2354
QY 311 PheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAla 330
DB 2355 ACGTTTGGAAAGCTGGCGAGATCTCCACAGAAAGCTGCGGTCAATGGCTTTTAAAGCA 2414
QY 331 MetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThr 350
DB 2415 ATGCTAAGACAGGACATGAGCTGTTTGTATGACCATAAAAACAGTACTGTCGTCATTTCT 2474
QY 351 ThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyVal 370
DB 2475 ACAAGACTTGGCACAGATGCTGCCCAAGTCCAAAGAGGCCACAGGAACCAAGTTGGCTTTA 2534
QY 371 LeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyr 390
DB 2535 ATTGACAGAAATATAGCTAACCTTGGAACTGGTATATCATATCATATCATATCATATCAT 2594
QY 391 GluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGlu 410
DB 2595 CAGTTAAACCTATTGCTATTAGCAGTTGTTCCAAATTTATTTGCTGTGTGTCAGGAATTTG 2654
QY 411 ThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisIleGlyLys 430
DB 2655 ATGAAATTTGTTGGCTGGAAATGCCAAAGAGATAAAAGAACTGGAAGCTGCTGGAAG 2714
QY 431 IleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAla 450
DB 2715 ATTGCAACAGAGCAATAGAAATATTAGGACAGTTGTGTCTTTGACCCAGGAAGAAAA 2774
QY 451 PheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAla 470
DB 2775 TTTGAATCAATGTATGTTGAAAAATTTGATGGACCTTACAGGAATTCGTGCAAGAGGCA 2834
QY 471 GlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAla 490
DB 2835 CACATCTATGGAATTTACTTTTAGTATCTCAAGAGCATTTATGATTTTCTTCTATGCCGT 2894
QY 491 GlyPheArgPheGlyValTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPhe 510
DB 2895 TGTTTTCGATTTGGTCATATCTCAATTGGAATGGACATATGCGCTTCAGAGATGTTAT 2954
QY 511 IleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAla 530
DB 2955 CTGTTGTTTCTGCAATTTGATTTGTCAGTGGCTCTAGGACATGCCAGTTCATTGCT 3014
QY 531 ProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGlyLysLys 550
DB 3015 CCAGACTATGCTAAAGCTGTGTCAGGCCCATTTATTCATGCTGTTTGAAGACAA 3074
QY 551 ProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeu 570
DB 3075 CCTCGATTGACAGCTACAGTGAAGAGGGGCTGAAGCCCTGATAAATTTGAAGAAATATA 3134
QY 571 GluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGly 590
DB 3135 ACATTTAATGAAGTCGTGTTCAACTATCCACCCGAGCAACCTGTCAGTGTCTCAGGGG 3194
QY 591 LeuSerLeuSerIleGluArgGlyThrValAlaPheValGlySerSerGlyCysGly 610
DB 3195 CTGAGCTCGAGGTGAAGAAAGGCCAGACACTAGCCCTGTCGTCAGCAGTGGCTGTGG 3254
QY 611 LysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlnValLeu 630
DB 3255 AAGAGCACGGTGTCTCAGCTCTCGAGCGGTCTACGACCCCTTGGCGGGGACAGTGTCT 3314

```


Qy	631	PheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAla	650
Db	3315	CTCGATGGTCAAGAACAAAGAACTCAATGTCTCAGTGGCTCAGAGCTCAACTCGGAATC	3374
Qy	651	ValProGlnGluProValLeuPheAsnCySerIleAlaGluAsnIleAlaTyGlyAsp	670
Db	3375	GTGTCTCAGAGAGCTATCCCTATTGTGATCGACATVTCGCGAGAATATTGGCTATTGGAGAC	3434
Qy	671	AsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHis	690
Db	3435	ACACGCCGGTGTGTATCACAGGATGAAATTGTAGTGCAGCCCAAGCTGCCCAACATACAT	3494
Qy	691	SerPheIleGluGlyLeuProGluLysTyfAsnThrGlnValGlyLeuLysGlyAlaGln	710
Db	3495	CCTTTCATCAGACGCTTACCCCAAAATATGAAACAAGAGTGGGAGATAAGCGGACTCAG	3554
Qy	711	LeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuGlnLysProLys	730
Db	3555	CTCTCAGAGGTCAAAACACAGAGATTGTATTGTGCCGAGCCCTCATCAGACAAACCTCAA	3614
Qy	731	IleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGln	750
Db	3615	ATCTCTCTGTGTGATGAGCTACATCAGCTCTGGATCTGGAAGTGAAGAAGTTGTCCAA	3674
Qy	751	HisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSer	770
Db	3675	GAAGCCCTGGACAAAGCCAGAGAAGCCGCACCTGCTATTGTGATTGCTCACGGCTGTGCC	3734
Qy	771	AlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleGlyGluGlnGly	790
Db	3735	ACCATCCAGAAATCAGACTTAATAGTGGTGTTCAGAAATGGGAGAGTCAAGGAGCATGGC	3794
Qy	791	ThrHisGlnGluLeuLeuArgAsnArgAspIleTyPheLysLeuValAsnAlaGlnSer	810
Db	3795	ACGATCAGCAGCTGCTGGCACAGAAGAGCATCTATTTTCATGTCAGTGTCCAGCT	3854
RESULT	14		
AX409652		3924 bp	linear
LOCUS	AX409652	Sequence 2299 from Patent WO0229103.	
DEFINITION	AX409652		
ACCESSION	AX409652.1	GI:21442357	
VERSION			
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	1		
AUTHORS	Alvares,C., Horne,D., Peres-da-Silva,S. and Vockley,J.G.		
TITLE	Gene expression profiles in liver cancer		
JOURNAL	Patent: WO 0229103-A 2299 11-APR-2002;		
FEATURES	GENE LOGIC INC (US)		
source	Location/Qualifiers		
	1..3924		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/notes="EMBL/GenBank Accession No. M23234"		
BASE COUNT	1145 a	790 c	977 g
ORIGIN		1012 t	
Alignment Scores:			
Pred. No.:	2,74e-175	Length:	3924
Score:	2465.50	Matches:	478
Percent Similarity:	76.46%	Conservative:	149
Best Local Similarity:	58.29%	Mismatches:	182
Query Match:	60.44%	Indels:	11
DB:	6	Gaps:	3
US-09-873-409-2	(1-812) x AX409652	(1-3924)	
Qy	2	ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyArgAspHisIleGlyVal	21

Db	1395	ATTGATGGCCAGAGATATTAGGAACCTTTAATGTAAACTATCTGAGGGAATCATTTGGTGTG	1455
Qy	22	VaiserGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysYrGlyArg	41
Db	1455	GTGAGTCAGGAGCGGTGCTGTTTTCCACCACCAATTTGTGAAATATTTTGTATGGCCGT	1514
Qy	42	AspAspValThrAspGluGluMetGluArgAlaIalAargGluAlaAsnAlaTyrAspPhe	61
Db	1515	GGAAATGTAACTGATGAGATAAAGAAGCTGTCAAGAGGGCCACGCCCTATGAGFTT	1574
Qy	62	IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer	81
Db	1575	ATCATGAAATACACAGAAATTTGCACCTTGTTGGAGAGAGAGGGCCAGCTGAGT	1634
Qy	82	GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu	101
Db	1635	GGTGGGCAGAAAGCAGAGGATCGCATTTGCACGTGCTGTTTGCACACCCCAAGATCCTT	1694
Qy	102	IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla	121
Db	1695	CTGCTGGATGAGGCCAGTCAGCATTTGGACACAGAAAGTGAAGCTGAGGTACAGCAGCT	1754
Qy	122	LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle	141
Db	1755	CTGGATGAAGCCAGAGAGCCGGACCACTTTGTATAGACACACCGACTGTCTACGGTC	1814
Qy	142	ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis	161
Db	1815	CGAAATGCAGATGTCATCGCTGGTGTGAGGATGAGTAATTTGTGGACGAAAGGACAC	1874
Qy	162	AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuVal	175
Db	1875	AGCGAATGATGAAGAAGGAGGAGGAGTACTTTCAAACTTGTCAACATCGACACATCAGGA	1934
Qy	176	-----MetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyr	192
Db	1935	AGCCAGATCCAGTCAGAAAGATTTTGAACTAAATGATGAAGAAGGCTGCCACTAGAAATGCC	1994
Qy	193	SerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLys---Ser	211
Db	1995	CCAAATGGCTGGAAATCTCGCCTATTATAGGCATTTCTACTCAGAAAAACCTTTAAAAATTC	2055
Qy	212	AspPheIleAspLysAla---GluGluSerThrGlnSerLysGluIleSerLeuProGlu	230
Db	2055	CAAAATGTCTCAGAAAGACCTTGATGTGAAACCCGATGGACTTTGAAGCAAAATGTGCCACCA	2114
Qy	231	ValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGly	250
Db	2115	GTGTCTCTTTCTGAAGGTCTCGAAACTTGAATAAACAGAAATGCCCCCTACTTTGTCTGGGA	2174
Qy	251	ThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLys	270
Db	2175	ACAGTATGGCATTTGCCAATGGGGGCTTCAGCGGCATTTTTTCAGTCATATTTCTCAG	2234
Qy	271	IleIleThrMetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyr	290
Db	2235	ATCATAGCGAATTTTGGACACAGCGCATGATGCAGTGAACGACGACAGAGTGCACATATTC	2294
Qy	291	SerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeu	310
Db	2295	TCTTTGATTTCTTATTCTTGGGAATATTCTTTTTTTTACTTTCTTCTTCAGGGTTTC	2355
Qy	311	PheTyrGlyArgAlaGlyIleLeuThrMetArgLeuArgHisLeuAlaPheLysAla	330
Db	2355	ACGTTTGGAAAGCTGCGAGATCCTCACCAGAAAGACTCGCGTCAATGGCTTTTAAAGCA	2414
Qy	331	MetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyClyLeuThr	350
Db	2415	ATGTAAAGACAGGACATGAGCTGGTTGTATGACCATTAATAAAACAGTACTTGTCATCTTCT	2474
Qy	351	ThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyVal	370
Db	2475	ACAAAGCTTGCACAGATGTCCCAAGTCCCAAGGAGCCACAGGAACCGAGTTGCTTTA	2534


```
QY 371 LeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyr 390
   ::: |||||
Db 2535 ATTGCACAGAAATAGTAACTTGGAACTGGTATTATCATATCATATCATATCATACGGTGG 2594

QY 391 GluMetThrPheIleIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGlu 410
   ::::: |||||
Db 2595 CAGTTAAACCTATTGCTATTAGCAGTGTTCAAATTTGCTGTGTGCAGAAATTTGTGAA 2654

QY 411 ThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLys 430
   ::: |||||
Db 2655 ATGAATTTGCTGGAAATGCTCAAAAGAGATAAAGAACTGGAAGCTGCTGGAAG 2714

QY 431 IleAlaThrGluAlaLeuGluAlaIleArgThrIleValSerLeuThrArgGlyLysAla 450
   |||||
Db 2715 ATTGCAACAGAGCAATAGAAATATTAGACAGTGTGTCTTTGACCAGGAAAGAAAA 2774

QY 451 PheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAla 470
   |||||
Db 2775 TTTGAATCAATGATGTTGAAAATTTGTATGGACCTTACAGGAATTTCTGTGCAGAGGCA 2834

QY 471 GlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAla 490
   |||||
Db 2835 CACATCTATGGAATTAATTTAGTATCTCACAGCAATTTATGTTATTTCTTATGCGCGT 2894

QY 491 GlyPheArgPheGlyAlaTyrIleuIleGlnAlaGlyArgMetThrProGluGlyMetPhe 510
   |||||
Db 2895 TGTTTTCGATTTGGTCATATCTCATTTGTAATGGACATATGCGCTTCAGAGATGTTATT 2954

QY 511 IleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAla 530
   |||||
Db 2955 CTGGTGTTCGCAATTTGTTTGTGTGAGTGGCTCTAGGACATGCCAGTTCATTGCTT 3014

QY 531 ProGluTyrSerLysAlaLysSerGlyAlaAlaHisPheAlaLeuLeuGlyLysLys 550
   |||||
Db 3015 CCAGACTATGCTAAAGCTAAGCTGTCTGAGCCCACTTATTCATGCTGTTTGAAGACAA 3074

QY 551 ProAsnIleAspSerArgSerGlnGluGlyLysProAspThrCysGluGlyAsnLeu 570
   |||||
Db 3075 CCTCTGATTGACAGCTACAGTGAAGAGGGGCTGAAGCCTGATAAATTTGAAGAAATATA 3134

QY 571 GluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGly 590
   |||||
Db 3135 ACATTTAATGAAGTGTGTTCACATCCACCCGAGCAACGTCAGTCTCAGGGG 3194

QY 591 LeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGly 610
   |||||
Db 3195 CTGAGCTGGAGGTGAAGAAAGCCAGACACTAGCCCTGTGGGCGACAGTGTGCTGGG 3254

QY 611 LysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeu 630
   |||||
Db 3255 AAGAGCAGGTGTCTCAGCTCCTGGAGCGGTCTACGACCCCTTGGCGGGACAGTGTCT 3314

QY 631 PheAspGlyValAspAlaLysGluLeuAsnValGlnThrPheArgSerGlnIleAlaIle 650
   |||||
Db 3315 CTCGATGGTCAAGAACCAAGAAACTCAATGTCCAGTGTGCTCAGAGCTCAACTCGGAATC 3374

QY 651 ValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAsp 670
   |||||
Db 3375 GTGTCTCAGAGCCTATCTATTGATCTGACGATTTGCCAGAAATATTCCTCTATGAGAC 3434

QY 671 AsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaIleHis 690
   |||||
Db 3435 AACAGCGGGTGTATCACAGGATGAAATTTGTGAGTGCAGCCGACCAAGCTGCCAACATACAT 3494

QY 691 SerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGln 710
   |||||
Db 3495 CTTTTCATCGACAGCTTATCCCCACAAATATGAACAAGAGTGGGAGATAAAGGGGCTCAG 3554

QY 711 LeuSerGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuGlnLysProLys 730
   |||||
Db 3555 CTCTCAGGAGGTCAAAAACAGAGGATGCTATTGTCGCCGAGCCCTCATACAGAACTCAA 3614
```

```
QY 731 IleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGln 750
   |||||
Db 3615 ATCTCCTCTGTTGGATGAAGCTACATCAGCTCTGGTACTGAAAGTCAAAAGGTTGTCCAA 3674

QY 751 HisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSer 770
   |||||
Db 3675 GAAGCCCTGGACAAGCCAGAGAGCCGACCTCATTTGATTGCTCACCGCTGCTCC 3734

QY 771 AlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGly 790
   |||||
Db 3735 ACCATCCAGATGCAGACTTAATAGTGTTCAGATGGAGAGTCAAGAGGATGGC 3794

QY 791 ThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 810
   |||||
Db 3795 ACGCATCAGCAGCTGCTGCGCACAGAAAGCATCTATTTTCAATGCTCAGTGTCCAGGCT 3854

RESULT 15
HUMMDR3
LOCUS Human membrane glycoprotein P (mdr3) mRNA, complete cds.
DEFINITION
ACCESSION M23234
VERSION M23234.1 GI:187501
KEYWORDS P-glycoprotein; membrane glycoprotein.
SOURCE Human liver, cDNA to mRNA, clone 3.27.
ORGANISM Homo sapiens
REFERENCE
AUTHORS van der Bliek, A.M., Koorman, P.M., Schneider, C. and Borst, P.
TITLE Sequence of mdr3 cDNA encoding a human P-glycoprotein
JOURNAL Gene 71 (2), 401-411 (1988)
MEDLINE 89138016
PUBMED 2906314
COMMENT Draft entry and computer-readable sequence for [1] kindly provided
          by P. Borst, 21-MAR-1989.
FEATURES
    source
        1..3924
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            <1..3924
            /product="mdr3 mRNA"
            33..3872
            /note="P-glycoprotein"
            /codon_start=1
            /protein_id="AAA36207.1"
            /db_xref="GI:307181"
            /translation="MDLEAAKNGCTAWPTSAEGDFELGISGKQKKTKTVMGTGLT
LFPSYDQDKLFWSLGTFIMALHAGSGPLMMIVFGEMTDKFDVDTAGNFSQVFNLSLSL
LNPKLILEEEMTRYAYYSGLGAGLVAAVYQVSWFTLAAGRQIRKIRKQFFHAILRQ
EIGFNDINDTTELNTRITDDISKTISEGIDKVGVMFFQAVATFFAGTIVGTFRGKWLTL
VIMASPIGLSAAWAKILSADFSDKELAAVAKAGAAVEALGAIRTVIAFGGQNKEL
ERYOKHLENAKEIGIKKAI SANISMGIAFLLLIVASYALAFWYSGTLVISEYTTIGNM
TVFSSILIGAFSVQQAAPCIDAFANARGAAVYVFDIINDNPKIDSFSERGHKPSIKG
NLEFNDVHFSYPSRANVKILKLNKLVQSGQTVVALVSSGCGKSTTVQLIQLRYPDDE
ATNYEFIMKLPQKFDLTVLGERGAQLSGQKQRIARALVRNPKILLDEATSKALVE
SEAEVQAALKAREGRTTVIAHRLSTVRNADVAGDEGVIVQSGSELMKKEGVY
FKLVNMTSGSOIQSEBEFELNDEKAAATRMAPNGWKSRLEFRHSOTKNLKSOMCOKSLD
VEDTGLEANVPVPYFLVKLNKTEWYFVVTGTCATANGGLQAPFVSIFSEIITAFG
PGDDAVAQKKNIFSLIFLIGITISFTFLQGTGFGKAGEILFRRRSWAKFKAWLKQ
DMSVDDHDKNSTGALSTRLATDAQVQGTGTRLATIAQNIANGTIGIISFIYGMQL
TLLLAIVPTIIVSGVIMKLAGNKRDKLEAAGKIATETATENTRTVSLHQRK
FESMVVKLYGPVRSVOKAIVGITFSIOAFMYFSYAGCFRGAFLIVNHNFRD
VILVFSALVFGAVALHASSFPADYAKLSAAHLFMLFRQPLIDSYSEGLKPDFK
EGNITFNEVYVNYPRANVPVLOGLSLEVKKGOTLALVSGCGSKGVQLLEPYDP
LAGTIVLGGQAKKLNQWLRAQIGVSPFILFDCISAIENAIYAGDSNRVVSQDEIVS
AAKAAITHPIETLPHKYEYRVDGKGTQLSGGQKQRIARALIRQPOILLDEATSA
LDTSEKVVQEQALDKAREGRTCTVIAHRLSTIQNADLIIVVFQNGRVKHEGTHQOLLAQ
KGIYFSMVSVQAGTQNL"
    BASE COUNT 1145 a 790 c 977 g 1012 t
    ORIGIN
```

Alignment Scores:

Pred. No.:	2.74e-175	Length:	3924
Score:	2465.50	Matches:	478
Percent Similarity:	76.46%	Conservative:	149
Best Local Similarity:	58.29%	Mismatches:	182
Query Match:	60.44%	Indels:	11
DB:	9	Gaps:	3

US-09-873-409-2 (1-812) x HUMMDR3 (1-3924)

Qy	2	ValAspGluAsnAspIleArgAlaLeuAsnValAArgHisTyrArgAspHisIleGlyVal	21
Db	1395	ATTGATGGCAGGATATTAGGAACCTTTATGTAACTATCTCAGGGAATCATTTGGTGTG	1454
Qy	22	ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg	41
Db	1455	GTGAGTCAGGAGCGGTGCTGTTTCCACCAAAATGCTGTAATAATTTGTTATGGCGGT	1514
Qy	42	AspAspValThrAspGluGluMetGluArgAlaAArgGluAlaAsnAlaTyrAspPhe	61
Db	1515	GGAAATGTAAACCATGGATGAGATAGAAGAAAGCTGTCAAGAGGCCCAACGCCCTATGAGTTT	1574
Qy	62	IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyValaGlnMetSer	81
Db	1575	ATCATGAATTAACACAGAAATTTGACACCTTGTTGGAGAGAGGGGCCCGCTGAGT	1634
Qy	82	GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu	101
Db	1635	GCTGGCAGAACACAGGATGCCATTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1694
Qy	102	IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaIle	121
Db	1695	CTGCTGGATGAGGCCACGTCAGCATTTGGACACAGAAAGTGAAGTGAAGTGAAGTGAAGT	1754
Qy	122	LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle	141
Db	1755	CTGGATAAGCCAGAGAGCGGACCCATTTGATGATGATGATGATGATGATGATGATGATGAT	1814
Qy	142	ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis	161
Db	1815	CGAAATGCAGATGTCATCGCTGGTTTGGAGTGGAGTAAATTTGGAGCAAGCAAGCCAC	1874
Qy	162	AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuVal-----	175
Db	1875	AGCGAATGATGAAGAAGGAAGGGGTGTTACTTCAAACTTTGTCAACATGCAGACATCAGGA	1934
Qy	176	-----MetSerGlnAspIleLysLysAlaAspGluGlnMetLysSerMetThrTyr	192
Db	1935	AGCCAGATCCAGTCAGAGAAATTTGAACATAATGAATGAAGAAAGCTGCCACTAGAAATGGCC	1994
Qy	193	SerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLys---Ser	211
Db	1995	CCAAATGGCTGGAAATCTCGCCTATTATTAGGCTTCTACTCAGAAAAACCTTAAATAATCA	2054
Qy	212	AspPheIleAspLysAla---GluGluSerThrGlnSerLysGluLysSerLeuProGlu	230
Db	2055	CAAAATGTGTGAGAGAGCCTTGATGTGGAACCGATGGACTTGAAGCAAAATGTGCCACCA	2114
Qy	231	ValSerLeuLeuLysIleLysLeuAsnLysProGluTyrProPheValValLeuGly	250
Db	2115	GTGTCCTTTCTGAAGTCTGGAACCTGAATAAAACAGAAATGGCCCTACTTTGTCGGGGA	2174
Qy	251	ThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLys	270
Db	2175	ACGATATGTCGCAATTCGCAATGGGGGCTTCAGCCGGCATTTTCATCATATCTCAGAG	2234
Qy	271	IleIleThrMetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyr	290
Db	2235	ATCATAGCATTTTGGACCGGCGATGATGTCAGTGAAGCAGCAGCAAGTGAACATATTC	2294
Qy	291	SerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeu	310
Db	2295	TCTTTGATTCTTCTTCTGGAATATTCTCTTTTACTTCTTCTCTCTCTCTCTCTCTCTCTCT	2354

Qy	311	PheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAla	330
Db	2355	ACGTTTGGGAAGCTGGCAGATCCTCACCAGAAGACTGCGGTCAATGGCTTTTAAAGCA	2414
Qy	331	MetLeuTyrGlnAspIleAlaTyrPheAspGluLysGluAsnSerThrGlyGlyLeuThr	350
Db	2415	ATGCTAGACAGGACATGAGCTGGTTGATGACCATTAACACAGTACTGTTGCACTTCT	2474
Qy	351	ThrLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyVal	370
Db	2475	ACAAAGCTTGCACAGATGCTGCCAAGTCCAGGAGCCACAGGAACAGGTTGGCTTTA	2534
Qy	371	LeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyr	390
Db	2535	ATTGCACAGAATATAGCTAACCTTGGAACTGGTATTATCATATCATTTATCTACGGTGG	2594
Qy	391	GluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGlu	410
Db	2595	CAGTTAACCTTATTTGCTATTAGCAGTTGTTCCATTATTGCTGTGTCAGGAATTTGTA	2654
Qy	411	ThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnLeuLysHisAlaGlyLys	430
Db	2655	ATGAAATTTGTTGGCTGGAAATGCCAAAAGAGATAAAAAGAACTGGAAGCTGCTGGAAG	2714
Qy	431	IleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAla	450
Db	2715	ATTGCAACAGAGGCAATAGAAAATATTAGGACAGCTTGTCTTTGACCCAGGAAGAAAA	2774
Qy	451	PheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAla	470
Db	2775	TTTGATCAATGATTTGTTGAAAATTTGATGACCTTACAGGAATTTCTGTCAGNAGGCA	2834
Qy	471	GlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaIle	490
Db	2835	CACATCTATGGAATTTACTTTTAGTATCTCACAGCATTTATGATTTTCTTCTATCCGGT	2894
Qy	491	GlyPheArgPheGlyAlaTyrIleIleGlnAlaGlyArgMetThrProGluGlyMetPhe	510
Db	2895	TGTTTTTCGATTTGGTGCATATCTCATTTGATGACATATGCGCTTCAGAGATGTTATT	2954
Qy	511	IleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAla	530
Db	2955	CTGGTGTTTTCTGCAATTTGTTGGTCAGTGGCTCTAGGACATGCCAGTTCATTGCT	3014
Qy	531	ProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLys	550
Db	3015	CCAGACTATGCTAAAGCTAAGCTGCTGCAGCCCATTTATTTCATGCTGTTTGAAGACAA	3074
Qy	551	ProAsnIleAspSerArgSerGlnGluLysLysProAspThrCysGluGlyAsnLeu	570
Db	3075	CCTCTGATTCAGCTACAGTCAAGAGGGGCTGAAGCCCTGATAAATTTGAAGGAATATA	3134
Qy	571	GluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGly	590
Db	3135	ACATTTAATTAAGTCGTGTTCAACTATCCACCAGCAAGAGCTGCCAGTCTTCAGGG	3194
Qy	591	LeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGly	610
Db	3195	CTGACCTTGGAGTGAAGAAAGCCAGACACTAGCCCTGGTGGGAGCAGTGGCTTGGG	3254
Qy	611	LysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeu	630
Db	3255	AAAGCAGCGGTGCTCAGCTCTCGAGCGGTTCTACAGACCCCTTGGCGGGGAGCAGTGT	3314
Qy	631	PheAspGlyValAspAlaLysGluLeuAsnValGlnTyrLeuArgSerGlnIleAlaIle	650
Db	3315	CTCGATGTCAAGAGCAAGAAACTCAATGTCCAGTGGCTCAGAGCTCAATCCGAATC	3374
Qy	651	ValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAsp	670
Db	3375	GTGCTCAGGAGCCTATCTCTATTGACATGTCAGCATTTGGAGAAATATTCCTATGAGAC	3434
Qy	671	AsnSerArgValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHis	690


```
Db 3435 AACAGCCGGGTTGTATCAGAGGATGAAATTTGTGAGTGCAGCCAAAGCTGCCAACATACAT 3494
QY 691 SerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGln 710
Db 3495 CCTTTCATCGAGAGCTTACCCCAATATGAAACAAGAGTGGGAGATAAGGGGACTCAG 3554
QY 711 LeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLys 730
Db 3555 CTCTCAGGAGGTCAAAAACAGAGGATTGCTATTGCCGAGCCCTCATCAGACACCTCAA 3614
QY 731 IleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGln 750
Db 3615 ATCCTCCTGTGGATGAAGCTACATCAGCTCTGGATACTGAAAGTGAAGAGGTGTCCAA 3674
QY 751 HisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSer 770
Db 3675 GAAGCCCTGGACAAAGCCAGAGAGCCGACCTGCTGTTGTGATTGCTCAGCCGCTGTCC 3734
QY 771 AlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGly 790
Db 3735 ACCATCCAGAAATGCAGACTTAATAGTGTGTTTCAGATGGGAGAGTCAAGGAGCATGCG 3794
QY 791 ThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheIysLeuValAsnAlaGlnSer 810
Db 3795 ACCGATCAGAGCTGCTGGCAGAGAAAGGCACTATTTTCAATGGTCAGTGTCCAGGCT 3854
```

Search completed: March 31, 2003, 02:09:55
Job time : 11128.4 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 30, 2003, 01:20:01 ; Search time 793.594 Seconds
(without alignments)
2304.231 Million cell updates/sec

Title: US-09-873-409-2

Perfect score: 4079

Sequence: 1 MVDENDIRALNVRHYRDHIG.....QELLRNRDIYFKLVNAQSVQ 812

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=xlh
-Q=/cgn2_1/USPRO_pool/US09873409/runat_27032003_115416_19216/app_query.fasta_1.7544
-DB=N_Geneseq_101002 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -DOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09873409 @CGN 1 1 4247 @runat_27032003_115416_19216 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_101002.*
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3919	96.1	3699	24	ABK83223 Human transporter
2	2465.5	60.4	3924	21	AAZ94742 Human ATP binding
3	2465.5	60.4	3924	21	AAZ88974 Human MDR-3 DNA.
4	2465.5	60.4	3924	24	ABN95801 Gene #2299 used to
5	2442	59.9	4279	22	AAD03488 Dog P-glycoprotein
6	2442	59.9	4279	22	AAD03504 Dog P-glycoprotein
7	2442	59.9	4279	22	AAD03505 Dog P-glycoprotein
8	2441	59.8	4186	22	AAF86127 Cynomolgous monke
9	2441	59.8	4195	22	AAF86128 Cynomolgous monke
10	2441	59.8	4279	22	AAD03506 Dog P-glycoprotein
11	2441	59.8	4369	21	AAZ52047 Rat multidrug resi
12	2441	59.8	4425	21	AAZ52048 Rat multidrug resi
13	2438	59.8	4317	22	AAD03489 Dog P-glycoprotein
14	2436	59.7	3840	24	ABL91687 Human polynucleoti
15	2436	59.7	3860	21	AAZ49332 Human wild-type mu
16	2436	59.7	3860	21	AAZ49333 Human G185V mutant
17	2436	59.7	3860	24	ABA94365 Human BCRP DNA rel
18	2436	59.7	3860	24	ABA94366 Human BCRP DNA rel
19	2436	59.7	3988	21	AAZ88973 Human MDR-1 DNA.
20	2436	59.7	4264	19	AAV66533 Mutated human P-gl
21	2436	59.7	4264	19	AAV66534 Mutated human P-gl
22	2436	59.7	4349	22	AAH57442 Human intestine ce
23	2436	59.7	4646	21	AAZ94738 Human ATP binding
24	2436	59.7	4646	24	AAD38994 Human mdr1 gene.
25	2436	59.7	4646	24	ABL68592 Kidney cancer rela
26	2436	59.7	4646	24	ABL68880 Kidney cancer rela
27	2436	59.7	4669	8	AAAT0752 Sequence of human
28	2436	59.7	6505	17	AAAT13394 Hybrid vector pSF-
29	2436	59.7	8630	21	AAZ24041 Retroviral M4 mdr-
30	2436	59.7	8630	21	AAZ24042 Retroviral vector
31	2433	59.6	4378	11	AAQ04522 Multidrug Resistan
32	2433	59.6	4669	14	AAQ52726 Sequence of human
33	2431.5	59.6	3912	24	ABK63653 Rat sequence of diffe
34	2425	59.5	4646	15	AAQ72872 Human multidrug re
35	2425	59.5	4788	21	AAZ49335 Murine multidrug r
36	2425	59.5	4788	24	ABA94368 Mouse BCRP DNA rel
37	2413.5	59.2	4233	21	AAZ90198 Rat mdr1b2 (multir
38	2413.5	59.2	4233	22	AAF27498 Rat mdr1b2 (multir
39	2412	59.1	4669	19	AAV32645 Human P glycoprote
40	2412	59.1	4669	24	ABK52041 cDNA encoding huma
41	2404.5	58.9	4189	21	AAZ49334 Murine multidrug r
42	2404.5	58.9	4189	21	AAZ49336 Murine multidrug r
43	2404.5	58.9	4313	14	AAQ38950 Mouse multidrug re
44	2372	58.2	4254	24	ABK63517 Rat sequence diffe
45	2280	55.9	2726	15	AAQ70907 Multidrug-resistan

ALIGNMENTS

RESULT 1	
ABK83223	
ID	ABK83223 standard; cDNA; 3699 BP.
XX	
AC	ABK83223;
XX	
DT	27-AUG-2002 (first entry)
XX	
DE	Human transporter and ion channel, TRICH14, Incyte ID 7472030CB1, cDNA.
XX	
KW	Human; ss; gene; transporter and ion channel; TRICH; transport disorder;
KW	neurological disorder; muscle disorder; immunological disorder; cancer;
KW	scleroderma; systemic lupus erythematosus; allergy; leukaemia;
KW	cell proliferative disorder; cervical cancer; breast cancer;
KW	neurodegenerative disorder; Parkinson's disease; Alzheimer's disease;
KW	myotonic dystrophy; catatonia; endocrine disorder; diabetes;
KW	Grave's disease; gastrointestinal disorder; Crohn's disease;
KW	renal disorder; Good pasture's syndrome; viral infection; cirrhosis;

KW bacterial infection; fungal infection; parasitic infection;
KW protozoal infection; helminthic infection; cardiovascular disorder;
KW atherosclerosis; hepatic disease.
XX Homo sapiens.
XX WO200240541-A2.
XX 23-MAY-2002.
XX 25-OCT-2001; 2001WO-US46055.
XX 27-OCT-2000; 2000US-243989P.
XX 03-NOV-2000; 2000US-245904P.
XX 09-NOV-2000; 2000US-247673P.
XX 17-NOV-2000; 2000US-249661P.
XX 20-NOV-2000; 2000US-252232P.
XX 01-DEC-2000; 2000US-250790P.
XX (INCY-) INCYTE GENOMICS INC.
XX PA
XX PI Tang YT, Yue H, Nguyen DB, Hafalia AJA, Elliott VS, Lu Y;
PI Walla NK, Yao MG, Baughn MR, Gandhi AP, Ding L, Sanjanwala M;
PI Ramkumar J, Arvizu C, Gietzen KU, Lal PG, Azimzai Y, Khan FA;
PI Thangavelu K, Thornton M, Lu DM, Tribouley CM, Warren BA;
PI Ison CH, Das D, Raumann BE, Policky JL, Kearney L;
XX WPI: 2002-463570/49.
XX P-PSDB; AB661544.
XX New transporters and ion channels (TRICH) polypeptides, useful for
XX diagnosing, preventing, and treating disorders associated with an
XX abnormal expression or activity of TRICH, e.g. immunological, muscular
XX or renal disorders -
XX Claim 5; Page 172; 178pp; English.
XX The invention relates to human transporters and ion channels (TRICH)
XX polypeptides, a naturally occurring amino acid sequence 90 % identical to
XX TRICH, a biologically active fragment of TRICH or an immunogenic fragment
XX of TRICH. Also included are an isolated polynucleotide encoding TRICH,
XX a recombinant polynucleotide comprising a promoter sequence operably
XX linked to the TRICH polynucleotide, a cell transformed with the
XX recombinant polynucleotide, a transgenic organism comprising the
XX recombinant polynucleotide, an isolated antibody that binds specifically
XX to TRICH, and screening for compounds which bind to TRICH, modulate
XX TRICH, modulate TRICH expression or are ant/agonists of TRICH.
XX The polypeptides are useful for diagnosing, treating, and
XX preventing transport, neurological, muscle, immunological disorders
XX (e.g. scleroderma, systemic lupus erythematosus, allergies), cell
XX proliferative disorders such as cancers (e.g. leukaemia, cervical or
XX breast cancers), neurodegenerative disorders (e.g. Parkinson's disease,
XX Alzheimer's disease), muscular disorders (e.g. myotonic dystrophy,
XX catatonias), endocrine disorders (e.g. diabetes, Grave's disease),
XX gastrointestinal disorders (e.g. Crohn's disease), renal disorders
XX (e.g. Good pasture's syndrome), viral, bacterial, fungal, parasitic,
XX protozoal and helminthic infections, cardiovascular disorders (e.g.
XX atherosclerosis), or hepatic diseases (e.g. cirrhosis) and many
XX other diseases and disorders detailed in the specification. They can also
XX be used in assessing the effects of exogenous compounds on the
XX expression of nucleic acid and amino acid sequences of transporters and
XX ion channels. TRICH or its fragments may also be used in screening for
XX compounds that specifically bind to and modulate the activity of TRICH.
XX The polynucleotides can be used to create knock-in humanised animals or
XX transgenic animals to model human disease. The present sequence
XX encodes a TRICH protein.

SQ Sequence 3699 BP; 1116 A; 707 C; 860 G; 1016 T; 0 other;

Alignment Scores:

Pred. No.: 0 Length: 3699
Score: 3919.00 Matches: 785
Percent Similarity: 96.80% Conservative: 1

Best Local Similarity: 96.67% Mismatches: 0
Query Match: 96.08% Indels: 26
DB: 24 Gaps: 1
US-09-873-409-2 (1-812) x ABK83223 (1-3699)
QY 1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
Db 1339 ATGGTGGATGAGATGACATCAGAGCTTTAAATGTGCGCATATTCGAGACCATTATGGA 1398
QY 21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleIleYrGly 40
Db 1399 GTGGTTAGTCAAGAGCCTGTTTTTGTTCGGACCCACCATCAGTAAACAATATCAAGTATGA 1458
QY 41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
Db 1459 CGAGATGATGTGACTGATCAAGAGATGGAGAGAGCAGCAGGGAAGCAATCGGTATGAT 1518
QY 61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80
Db 1519 TTTATCATGGAGTTTCTTAATAAATTTAATACATTTGTTAGGGGAAAAAGGAGCTCAATG 1578
QY 81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProIle 100
Db 1579 AGTGAGGGGAGAAACAGAGGATCGCAATTCGTGCTTGTAGTTTCGAACCCCAAGATT 1638
QY 101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120
Db 1639 CTGATTTTAGATGAGGCTACGTCTGCCCTGGATTGAAAGCAAGTCAGCTGTTCAAGCT 1698
QY 121 AlaLeuGluLysAlaSerLysGlyArgThrThrIleValAlaAlaHisArgLeuSerThr 140
Db 1699 GCACCTGGAGAGGCGAGCAAGGTCGAGCTACAATCGTGGTAGCACCCGACTTTCTACT 1758
QY 141 IleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAla 160
Db 1759 APTCGAAGTGCAGATTGTTGTTGACCTTAAGGATGGAATGCTCGCGGAGAAAGAGCG 1818
QY 161 HisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIle 180
Db 1819 CATGCTGAACTAATGGCAAAAGAGGCTCTATATTTACTTGTGATGTCACAGGATATT 1878
QY 181 LysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSer 200
Db 1879 AAAAAGCTGATGAACAGATGAGTCAATGACATATTTCTCTGAAAGAAAGAACCACTCA 1938
QY 201 LeuProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSer 220
Db 1939 CTTCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTCATTGACAAGGCTGAGGAATCC 1998
QY 221 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 240
Db 1999 ACCCAATCTAAGAGATGAAGTCTTCTGAAAGTCTCTCTATTAAAAATTTTAAAGTTAAAC 2058
QY 241 LysProGluTyrProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 260
Db 2059 AAGCTGATGAGCCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2118
QY 261 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLys 280
Db 2119 CATCAGATATTTTCCATCATCTTTTGCACAAAAATTTATAACCATGTTTGGAAATAATGATAA 2178
QY 281 ThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleIleGlyValIle 300
Db 2179 ACCAATTAAGCATGATGCAGAAATTTATTCATGATATTCGTATTCGTATTCGTATTCGTAT 2238
QY 301 CysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThr 320
Db 2239 TGTCTTGTGCTGATTTTTCATG----- 2259
QY 321 MetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAsp 340
Db 2260 -----CAGGATATTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2280

Qy	701	AsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAla	720
Dd	3361	AACACACAAGTTGGACTGAAGAGGCACACAGCTTTCTGGCGCCAGAAACAAAGACTAGCT	3420
Qy	721	IleAlaARGalalaLeuGlnLysProLysIleLeuLeuLeuASPgluaIathrSerAla	740
Dd	3421	ATTGCCAAGGGCTCTTCTCAAAAACCCAAATTTTATTGTGTGGATGAGCCACTTCAGCC	3480
Qy	741	LeuASpAsnASerGlnLysValValGlnHisAlaLeuASpLYsaLaArGThrGlyArg	760
Dd	3481	CTCGATAATCACACAGTCAGAAGCTGGTTCAGCATGCGCTTGATAAAGCCAGACGGGAAGG	3540
Qy	761	ThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaASpLeuValVal	780
Dd	3541	ACATGCTTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAACGCGAGATTGATAGTGTT	3600
Qy	781	LeuHiEAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgasp	800
Dd	3601	CTGCACAATGGAAGATAAAGGAACAAGAACTCATCAAGAGCTCTCAGAAATCGAGAC	3660
Qy	801	IleTyrrPheLysLeuValAsnAlaGlnSerValGln	812
Dd	3661	ATATATTTTAAAGTTAGTGAATGACAGTCAGTGCAG	3696
RESULT 2			
ID	AAZ94742 standard; cDNA; 3924 BP.		
XX			
AC	AAZ94742;		
XX			
DT	01-AUG-2000 (first entry)		
XX			
DE	Human ATP binding cassette ABCB4 (MDR3) CDNA.		
XX			
KW	ABCB4; ATP binding cassette; human; cholesterol; lipid disorder; attherosclerosis; lipid disorder; dyslipidemia; psoriasis;		
KW	lupus erythematosus; diagnosis; gene therapy; MDR3;		
KX	multidrug resistance; chromosome 7q21; ss.		
OS	Homo sapiens.		
XX			
FN	WO200018912-A2.		
XX			
PD	06-APR-2000.		
XX			
PF	21-SEP-1999; 99WO-BP06991.		
XX			
PR	25-SEP-1998; 98US-0101706.		
XX			
PA	(FARB) BAYER AG.		
XX			
F1	Schmitz G, Klucken J;		
XX			
DR	WPI; 2000-293151/25.		
XX			
PT	Adenosine triphosphate binding proteins useful for identifying agents		
PT	for treating atherosclerosis and other inflammatory disorders -		
XX			
PS	Claim 9; Page 117-119; 154pp; English.		
XX			
CC	The present sequence is that of human ATP binding cassette		
CC	subfamily B protein ABCB4 cDNA. The cDNA was identified using a		
CC	differential display method in which monocytes from peripheral		
CC	blood were subjected to macrophage differentiation and cholesterol		
CC	loading with acetylated low density lipoproteins and subsequent		
CC	deloading with high density lipoprotein (HDL3) to identify		
CC	cholesterol sensitive genes. The gene maps to chromosome 7q21		
CC	and is also termed MDR3 (multidrug resistance). The invention		
CC	provides cholesterol-sensitive ABC genes (see AA294734-63). These		
CC	genes, and polypeptides encoded by them, can be used for diagnostic		
CC	and therapeutic applications, and for biochemical or cell-based		
CC	assays to screen for pharmacologically active modulator compounds		

CC other inflammatory diseases such as psoriasis and lupus
CC erythematousus.

SQ Sequence 3924 BP; 1145 A; 790 C; 977 G; 1012 T; 0 other;

Alignment Scores:

Pred. No.: 7,766-229 Length: 3924
Score: 2465.50 Matches: 478
Percent Similarity: 76.46% Conservative: 149
Best Local Similarity: 58.29% Mismatches: 182
Query Match: 60.44% Indels: 11
DB: 21 Gaps: 3

US-09-873-409-2 (1-812) x AA294742 (1-3924)

QY 2 ValAspGluAsnAspIleArgAlaLeuAenValArgHisTyrArgAspHisIleGlyVal 21
DB 1395 ATTGATGGCGAGGATATTAGGAACCTTTAATGTAACATATCTCAGGGAATCATTTGGTGG 1454
QY 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
DB 1455 GTGAGTCAGGAGCGGTGCTGTTTCCACCAACATTTGCTGAAATATTTGTTATGGCGGT 1514
QY 42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
DB 1515 GGAAATGTAACCATGATGAGATAAAGAAAGCTGTCAAAGAGCGCCAAACGCTATGAGTTT 1574
QY 62 IleMetGluPheProAsnLysPheAenThrLeuValGlyGluLysGlyAlaGlnMetSer 81
DB 1575 ATCATGAAATACCACAGAAATTTTGACCCCTGGTTGGAGAGGGGGCCCGCTGAGT 1634
QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
DB 1635 GGTGGCCAGACAGAGGATCGCCATTGACGTCGCTGTTCCGCAACCCCAAGATCCTT 1694
QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
DB 1695 CTGCTGGATGAGGCCACGTCAGCATTTGGACACAGAAAGTGAAGCTGAGGTACAGGCAGCT 1754
QY 122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
DB 1755 CTGGATAAGCCAGAGAGCGGACGACCATTTGTATGACACACCGCATGTCTACGGTC 1814
QY 142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
DB 1815 CGAAATGCAGATGTCATCGCTGGTGTGGAGTGGAGTAATTTGGGAGCAAGAGCCAC 1874
QY 162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuVal----- 175
DB 1875 AGCGAACTGATGAAGAAGAAAGGGGTGACTTCAAACTTGTCAACATGCGAGACATCAGGA 1934
QY 176 -----MetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyr 192
DB 1935 AGCCAGATCCAGTCAGAAAGATTTGAACATAATGATGAAGAGCTGCCACTAGATGGCC 1994
QY 193 SerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLys---Ser 211
DB 1995 CCAAAATGGCTGGAATCTCGCCTATTATAGGCATTCTACTCAGAAAAACCTTAAATATCA 2054
QY 212 AspPheIleAspLysAla---GluGluSerThrGlnSerLysGluIleSerLeuProGlu 230
DB 2055 CAAATGTGTGAGAAGAGCCCTTGATGTGGAAACCGATGAGCTGTAAGCAAAATGTGCCACCA 2114
QY 231 ValSerLeuLeuLysIleLeuLysLeuAenLysProGluTyrProPheValValLeuGly 250
DB 2115 GTGTCCTTTCTGAAGTCTCTGAAACCTGAATAAACAGATAGGCCCTACTTTGTCTGGGA 2174
QY 251 ThrLeuAlaSerValLeuAenGlyThrValHisProValPheSerIleIlePheAlaLys 270
DB 2175 ACAGATATGTCATTCGCAATGGGGGCTTCAGCCGCGCATTTTTCAGTCATATTTCTCAGAG 2234
QY 271 IleIleThrMetPheGlyAsnAenAspLysThrThrLeuLysHisAspAlaGluIleTyr 290
DB 1935 CTGCGATGTTCAAGAGAGCAAGAAACTCAATGTCCAGTGGCTCAGAGCTCACTCGGAATC 3374

DB 2235 ATCATAGCGATTTTTGGACCGGCGATGATGAGTGAAGACGACGAGAAAGTGCAACATATTC 2294
QY 291 SerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeu 310
DB 2295 TCTTTGATTTCTTATTTCTGGAAATTTATTTCTTTTCTTTTCTCTCTCTCAGGTTTC 2354
QY 311 PheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAla 330
DB 2355 ACGTTTGGAAAGCTGGCGAGATCTCCACAGAAAGACTCGCGTCAATGGCTTTTAAAGCA 2414
QY 331 MetLeuTyrGlnAspIleAlaTyrPheAspGluLysGluAenSerThrGlyGlyLeuThr 350
DB 2415 ATGCTAAGACAGGACATGAGCTGGTTTATGACCAATAAAACAGTACTGGTGCATTTCT 2474
QY 351 ThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyVal 370
DB 2475 ACAAGACTTGCACAGATGCTGCCCAAGTCCCAAGGACGACACAGGAACAGGTGTGCTT 2534
QY 371 LeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyr 390
DB 2535 ATTGCACAGAATATAGCTAAACCTTGAACTGGTATATATCATATCATTTATCTACGGTTGG 2594
QY 391 GluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGlu 410
DB 2595 CAGTTAAACCTTATTTCTATAGCAGTTGTTCCAAATTTATTTGCTGTGAGGAATTTGTA 2654
QY 411 ThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLys 430
DB 2655 ATGAAATTTGTTGGCTGGAATGCCCAAGAGATAAAGAACTGGAAGCTCTGGAAG 2714
QY 431 IleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAla 450
DB 2715 ATTGCAACAGAGGCAATAGAAAATATTAGGACAGTTGTCTTTGACCCAGGAAAGAAA 2774
QY 451 PheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAla 470
DB 2775 TTTGAATCAATGTATGTTGAAATTTATGAGACCTTACAGGAATCTGTGCGAAGGCA 2834
QY 471 GlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAla 490
DB 2835 CACATCTATGGAATTTACTTTTAGTATCTCACAAGCATTTATGTATTTTCTTATCCGCT 2894
QY 491 GlyPheArgPheGlyAlaTyrIleuIleGlnAlaGlyArgMetThrProGluGlyMetPhe 510
DB 2895 TGTTTTTCGATTTGTTGCATATCTCATTTGTAATGACATATATCGCTTTCAGAGATGTTAT 2954
QY 511 IleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAla 530
DB 2955 CTGGTGTCTTCTGCAATTTGTTGTCAGTGGCTCTAGGACATGCCAGTTCATTGCT 3014
QY 531 ProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLys 550
DB 3015 CCAGACTATGCTAAAGCTAAGCTGTCTGCAGCCCATTTATTCATGCTGTTTGAAGACAA 3074
QY 551 ProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeu 570
DB 3075 CCTCTGATTCAGCTACAGTGAAGAGGGCTGAAGCCCTGATAAAATTTGAAGGAAATATA 3134
QY 571 GluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGly 590
DB 3135 ACATTTAATGAAGTCGTGTTCAACTATCCACCCGAGCAACGTCGCTGCTTCAGGGG 3194
QY 591 LeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGly 610
DB 3195 CTGAGCCTGGAGTGAAGAAAGCCAGACACTAGCCCTGGTGGGCGAGCTGGCTGGG 3254
QY 611 LysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeu 630
DB 3255 AAGAGACCGTGTGTCAGCTCTCTGAGCGGTTCTACGACCCCTTGGCGGGGACAGTGCTT 3314
QY 631 PheAspGlyValAspAlaLysGluLeuAenValGlnTyrLeuArgSerGlnIleAlaIle 650
DB 3315 CTCGATGTTCAAGAGAGCAAGAACTCAATGTCCAGTGGCTCAGAGCTCACTCGGAATC 3374

QY 193 SerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLys---Ser 211
DB 1995 CCAAAATGGCTGGAAATCTCGCCCTATTATTAGGCATTTCTACTCAGAAACAACTTAAATAATCA 2054
QY 212 AspPheIleAspLysAla---GluGluSerThrGlnSerLysGluIleSerLeuProGlu 230
DB 2055 CAAATGTGTGAGAGAGCTTGTATGTGGAAACCGATGGACTTGAACCAATGTGCCACCA 2114
QY 231 ValSerLeuLysLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGly 250
DB 2115 GTGTCTCTTCTGAAGTCTGAACTGAATAAAACAGAAATGGCCCTACTTTGTGTGGGA 2174
QY 251 ThrLeuAlaSerValLeuAsnGlnThrValHisProValPheSerIleIlePheAlaLys 270
DB 2175 ACAGTATGTGCCATTCGCCAATGGGGGCTTCAGCCGGCATTTTTCAGTCATATTTCTCAGAG 2234
QY 271 IleIleThrMetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyr 290
DB 2235 ATCATAGCGATTTTGGACCGCGATGTATGTAGTGAACGAGAGAGTGCACATATTTTC 2294
QY 291 SerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeu 310
DB 2295 TCTTTCATTTCTTATTCTGGGAATATTCTTTTCTTTTACTTTCTTCTCCTCAGGGTTTC 2354
QY 311 PheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAla 330
DB 2355 ACGTTTGGGAAAGCTGGCAGATCTCCACAGAAAGACTCGCGTCAATGGCTTTTAAAGCA 2414
QY 331 MetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThr 350
DB 2415 ATGCTAAGACAGACATGAGCTGGTTGTGTGATGACCAATAAAACAGTACTGCTGCACITTTCT 2474
QY 351 ThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyVal 370
DB 2475 ACAAGACTTGGCCACAGATGCTGCCAAGTCCAAGGACCAAGGACCAAGTGGCTTTA 2534
QY 371 LeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrp 390
DB 2535 ATTGCAACAGAAATATAGCTAACCTTGAACCTGGTATTTATCATATCATTTATCTACGGTTGG 2594
QY 391 GluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGlu 410
DB 2595 CAGTTAACCTTATGCTATTAGCAGTTGTTCATTAATTATTGCTGTGTCAGAAATTTGTGAA 2654
QY 411 ThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLys 430
DB 2655 ATGAATTTGTGCTGGAATGCGCAAGAGATATAAAGAACTGGAAGCTGCTGGAAAG 2714
QY 431 IleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAla 450
DB 2715 ATTGCAACAGAGCAATAGAAATATATTAGACAGTTGTGTCTTTGACCCAGGAAAGAAA 2774
QY 451 PheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAla 470
DB 2775 TTGGAATCAATGATTTGAAATAATTGTATGACCTTACAGCAATTTCTGTGCAGAGGCA 2834
QY 471 GlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAla 490
DB 2835 CACATCTATGGAATTTACTTTTAGTATCTCACAAGCATTTATGATTTTCTATGCCGT 2894
QY 491 GlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPhe 510
DB 2895 TGTGTTTCGATTTTGGTGCATATCTCATTTGTAATGGACATATGGCTTCAGAGATGTTATT 2954
QY 511 IleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAla 530
DB 2955 CTGGTGTCTTCTGCAATTTGATTTGTGTGAGTGGCTCTAGGACATGCCAGTTCATTGCT 3014
QY 531 ProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuGluLysLys 550
DB 3015 CCAGACTATGCTAAAGCTGTGTCAGCCCACTTATTCATGCTGTTTGAAGACAA 3074
QY 551 ProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeu 570

DB 3075 CCTCTGATTGACAGCTTACAGTGAAGAGGGGTGAAGCCCTGATAAATTTTGAAGAAATATA 3134
QY 571 GluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGly 590
DB 3135 ACATTTAATGAATCGTGTTCACATATCCACCCGAGCAACAGTCCAGTGTCTCAGGGG 3194
QY 591 LeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGly 610
DB 3195 CTGAGCTGGAGTGAAGAAAGCCAGACACTAGCCCTGGTGGCAGCAGTGCCTGTGGG 3254
QY 611 LysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeu 630
DB 3255 AAGACACCGTGGTCCAGCTCTCGAGCGGTTCACGACCCCTTGGCGGGGACAGTGCTT 3314
QY 631 PheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIle 650
DB 3315 CTCGATGTCAAGAACCAAGAACTCAATGTCCAGTGGCTCAGAGCTCACTCGAATC 3374
QY 651 ValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAsp 670
DB 3375 GTGTCTCAGGAGCTTATCTATTGACTGCAGCATTTGCCAGAAATATTGCTTATGGAGAC 3434
QY 671 AsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAsnIleHis 690
DB 3435 AACACCCCGGTGTATCACAGATGAAATTTGTAGTGCAGCCCAAGCTGCCAACATACAT 3494
QY 691 SerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGln 710
DB 3495 CTTTTCATCGAGACCTTATCCCAACCAATATGAAACAGAGTGGGAGATAAGGGACTCAG 3554
QY 711 LeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuGlnLysProLys 730
DB 3555 CTCTCAGGAGTCAAAAACAGAGGATGCTATTGCTGCCAGCCCTCATCAGACAACTCAA 3614
QY 731 IleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGln 750
DB 3615 ATCTCCCTCTTGGATGAAGCTACATCAGCTCTGGATCTGGAAGTGAAGAAAGTTGTCCAA 3674
QY 751 HisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSer 770
DB 3675 GNAGCCCTGGCAAGCCAGAGAGCCGACCTGCTCATTTGATTGCTCACCCCTGTCC 3734
QY 771 AlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGly 790
DB 3735 ACCATCCAGAAATGCAGACTTAATAGTGTGTTTCAGAAATGGAGAGTCAAGGAGCATGC 3794
QY 791 ThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 810
DB 3795 ACGCATCAGCAGCTCTGCGCAGAAAGGCATCTATTTTCAATGTCAGTGTCCAGGCT 3854
RESULT 4
ABN95801
ID ABN95801 standard; DNA; 3924 BP.
XX
AC ABN95801;
XX
DT 13-AUG-2002 (first entry)
XX
DE Gene #2299 used to diagnose liver cancer.
XX
Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
metastatic liver tumour; cytostatic; expression profile; disease state;
disease progression; drug toxicity; drug efficacy; drug metabolism.
XX
OS Homo sapiens.
XX
FN WO200229103-A2.
XX
PD 11-APR-2002.
XX
PF 02-OCT-2001; 2001WO-US30589.
XX

PR 02-OCT-2000; 2000US-237054P.

XX (GENE-) GENE LOGIC INC.

XX Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;

XX WPI; 2002-426119/45.

XX Diagnosing and detecting the progression of liver cancer,

PT hepatocellular carcinoma or metastatic liver tumor in a patient,

PT involves detecting the level of expression of two or more genes in a

XX liver tissue sample

XX Claim 1; SEQ ID NO 2299; 298bp; English.

XX The invention relates to a novel method for diagnosing and detecting the
CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
CC tumour in a patient, and differentiating metastatic liver cancer from
CC hepatocellular carcinoma in a patient, involving detecting the level of
CC expression of two or more genes represented in ABN93503-ABN97455 in a
CC tissue sample. The method of the invention has hepatotropic, and
CC cytostatic activity. The method is useful for diagnosing and detecting
CC the progression of liver cancer, hepatocellular carcinoma and metastatic
CC liver carcinoma in a patient. The method is useful for identifying
CC expression profiles which serve as useful diagnostic markers as well as
CC markers that can be used to monitor disease states, disease progression,
CC drug toxicity, drug efficacy and drug metabolism.

CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 3924 BP; 1145 A; 790 C; 977 G; 1012 T; 0 other;

Alignment Scores:

Pred. No.:	7,76e-229	Length:	3924
Score:	2465.50	Matches:	478
Percent Similarity:	76.46%	Conservative:	149
Best Local Similarity:	58.29%	Mismatches:	182
Query Match:	60.44%	Indels:	11
DB:	24	Gaps:	3

US-09-873-409-2 (1-812) x ABN95801 (1-3924)

QY 2 ValAspGluAsnAspIleArgAlaLeuAenValArgHisTyArgAspHisIleGlyVal 21

DB 1395 AITGATGGCCAGATATTAGGAATTTAATGTAACCTATCTGAGGGAATCATTTGGTGTG 1454

QY 22 ValSerGlnProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyArg 41

DB 1455 GTGAGTCAGGAGCGGTGCTGTTTCCACCACCAATTTGCTGAAATATTTGTTATGCCCT 1514

QY 42 AspAspValThrAspGluMetGluArgAlaAalArgGluAlaAsnAlaTyArgPhe 61

DB 1515 GGAATGTAAACCATGATGAGATAAAGAGCTGTCAAGAGGCCAACCCCTATGAGTTT 1574

QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValcylGluLysGlyAlaGlnMetSer 81

DB 1575 ATCATGAATTAACCCAGAGAAATTTGACCCCTGTTTGGAGAGAGGCGCCAGCTGAGT 1634

QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101

DB 1635 GGTGGGACAGAGCAGAGGATCGCATTTGCAGCTGCCCTGTTTCCGCAACCCCAAGATCCTT 1694

QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121

DB 1695 CTGCTGGATGAGCCACGTCAGCATTTGGACACACAGAAAGTGAAGCTGAGTACAGGCAGCT 1754

QY 122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141

DB 1755 CTGGATAGGCCAGAGAGCCGACACCATTTGTGATAGCACACCCGACTGTCTACGGTCT 1814

QY 142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161

DB 142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161

DB	1815	CGAATGCAGATGTCATCGCTGGGTTTCAGGATGGAGTAATTTGTGGCAAGGAAGCCAC	1874
QY	162	AlaGluLeuMetAlaLysArgGlyLeuTyTySerLeuVal	175
DB	1875	AGCGAACTGTGAAGAGGAAGGGGTGTACTTTCAAACTTTGTCAACATGCAGACATCAGGA	1934
QY	176	-----MetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTy	192
DB	1935	AGCCAGATCCAGTCCAGAAGAAATTTGAACATAATGATGAAAAGGTCGCCACTAGATGGCC	1994
QY	193	SerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLys--Ser	211
DB	1995	CCAAATGGCTGGAATCTCGCTATTAGGCATTTACTACAGAAAACCTTAAAAATTC	2054
QY	212	AspPheIleAspLysAla---GluGluSerThrGlnSerLysGluIleSerLeuProGlu	230
DB	2055	CAATGTGTGCAAGAGCGCTTGATGTGAAACCGATGACTTGAAGCAAAATGTGCCACCA	2114
QY	231	ValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTyTrpPropheValValLeuGly	250
DB	2115	GTGTCTTTCTGAAGGTCTGAACTGAATAAAACAGAAATGGCCCTACTTTGTCTGGGA	2174
QY	251	ThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLys	270
DB	2175	ACAGTATGTGCTCCATTCGCAATGGGGGCTTCAGCCGCGATTTTCATCATATTCACAG	2234
QY	271	IleIleThrMetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluLeTy	290
DB	2235	ATCATAGCGATTTTGGACCGCGCATGATGCAGTGAAGCAGCAGCAAGTGCACATATTC	2294
QY	291	SerMetIlePheValIleLeuGlyValIleCysPheValSerTyPheMetGlnGlyLeu	310
DB	2295	TCTTTGATTTTCTTATTTCTGGGAATTTCTTTTCTTTCTTTCTTTCTTTCTTTCTTT	2354
QY	311	PheTyArgArgAlaGlyGluLeuThrMetArgLeuArgHisLeuAlaPheLysAla	330
DB	2355	ACGTTTGGGAAGCTGGCGAGATCTCCACAGAACAGTCGGCTCAATGGCTTTTAAAGCA	2414
QY	331	MetLeuTyGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThr	350
DB	2415	ATGCTAAGACAGGACATGAGCTGTTTGTATGACCATAAAACAGTACTGTTCTTCT	2474
QY	351	ThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyVal	370
DB	2475	ACAAGACTTGCACAGATGCTGCCCAAGTCCAAGGAGGCACAGGAACAGGTGGCTTTA	2534
QY	371	LeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyGlyTy	390
DB	2535	ATTGCACAGAAATATAGCTTAACCTTGGAACTGGTATTTATCATATCATTTATCTAC	2594
QY	391	GluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGlu	410
DB	2595	CAGTTTAAACCTTATGCTATTAGCAGTGTTCNATTTATGCTGTCTGAGAAATTTGTA	2654
QY	411	ThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGlyLeuLysHisAlaGlyLys	430
DB	2655	ATGAAATGTTTGGCTGGAAATGCCAAAAGAGATAAAAGAACTGGAAGCTGCTGGAAG	2714
QY	431	IleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAla	450
DB	2715	ATTGCAACAGAGGCAATAGAAAATATTAGGCAGTGTGTCTTTTACCCAGGAAAGAAA	2774
QY	451	PheGluGlnMetTyGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAla	470
DB	2775	TTTGAATCATGTATGTTGAAAATTTGATGGACCTTACAGGAATTTGTGCAAGAGGCA	2834
QY	471	GlnIleIleGlySerCysTyzAlaPheSerHisAlaPheIleTyPheAlaTyAlaAla	490
DB	2835	CACATCTATGGAATTTACTTTTAGTATCTCACAAGCATTTATGTATTTTCTATGCGGT	2894
QY	491	GlyPheArgPheGlyAlaTyIleuIleGlnAlaGlyArgMetThrProGluGlyMetPhe	510
DB	2895	TGTTTTCGATTTGGTGCATATCTCATTTGTGAATGGACATATCGCTTTCAGAGATGTT	2954


```
Qy 511 IleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAla 530
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2955 CTGGTGTTCGCAATGTAATTTGGTGCAGTGGCTTAGACATCCAGTTCATTGCT 3014
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 531 ProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuLysLys 550
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3015 CCAGACTATGCTAAAGCTAGCTGTCTGCAGGCCACTTATTCATGCTGTTGAAAGACAA 3074
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 551 ProAsnIleAspSerArgSerGlnGluGlyLysLeuPheAspThrCysGluGlyAsnLeu 570
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3075 CCTCTGATTGACAGCTACAGTGAAGAGGGGCTGAAGCCCTGATAAATTTGAAGAAATATA 3134
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 571 GluPheArgGluValSerPheTyrProCysArgProAspValPheIleLeuArgGly 590
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3135 ACATTTAATGAAGTGTGTTCAACTATCCACCAGCAACACGTCAGTGTTCAGGGG 3194
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 591 LeuSerIleuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGly 610
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3195 CTGAGCCTGGAGGTGAAGAAAGGCCAGCACTAGCCCTGTGGCGCAGTGGCTGTGGG 3254
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 611 LysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyValLeu 630
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3255 AAGACACGGTGTTCAGCTCTGGAGCGGTTCTAGACCCCTTGGCGGGGACAGTCTT 3314
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 631 PheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIle 650
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3315 CTGATGGTCAAGCAAGCAAGAACTCAATGTCTCAGTGGCTCAGAGCTCACTCGGAATC 3374
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 651 ValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAsp 670
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3375 GTCTCTCAGAGGCTATCTCTATTGATGAGTGAATTTGTGAGTGCAGCCAAAGCTGCACATACAT 3434
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 671 AsnSerArgValValProLeuAspGluIleLysGluAlaAsnAlaAsnIleHis 690
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3435 AACAGCCGGGTTGTATCAGAGTGAATTTGTGAGTGCAGCCAAAGCTGCACATACAT 3494
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 691 SerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLysGlyAlaGln 710
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3495 CCTTTCATCAGACGTTACCCCAACATATGAACACAGAGTGGAGATAGGGGACTCAG 3554
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 711 LeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuLysProLys 730
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3555 CTCTCAGGAGGTCAAAAACAGAGGATTTGTTATTCGCCGAGCCCTCATCAGAACCTCAA 3614
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 731 IleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValGln 750
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3615 ATCTCTCTGTGGATGAAGCTACATCAGCTCTGGATCTGAAGTGAAGAGGTGTGCCAA 3674
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 751 HisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSer 770
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3675 GAAGCCCTGGCAAAAGCCAGAGAGCCGACCTGTCATTGTGATTCCTACCCGCTGTCC 3734
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 771 AlaIleGluAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGly 790
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3735 ACCATCCGAATGCAGACTTAATAGTGTGTTTCAGAAATGGGAGAGTCAAGGAGCATGCG 3794
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 791 ThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 810
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3795 ACGCATCAGCAGCTGTGGCACAGAAAGGCACTATTTTCAATGTGTCAGTGTCCAGGCT 3854
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
RESULT 5
AAD03488
ID AAD03488 standard; cDNA; 4279 BP.
XX
AC AAD03488;
XX
DT 13-JUN-2001 (first entry)
DE
XX Dog P-glycoprotein (PGP) cDNA #1.
XX
KW Dog; P-glycoprotein; PGP; multidrug transporter; MDR1;
drug bioavailability; transgenic animal; genetic model; ss.
```

```
XX
OS Canis familiaris.
FH Key Location/Qualifiers
FT CDS 17..3862
/*tag= a
/*product= "Dog P-glycoprotein (PGP) #1"
PN WO200123540-A2.
XX
PD 05-APR-2001.
PF 28-SEP-2000; 2000WO-US26767.
PR 28-SEP-1999; 99US-0156510.
XX (GENT-) GENTEST CORP.
PA
XX Stocker PU, Steimel-crespi DT, Crespi CL, Reif TC, Patten CU;
PI P-PSDB; AAE00303.
DR WPI; 2001-235373/24.
XX
PS Claim 3; Page 58-63; 111pp; English.
CC
CC The invention relates to dog P-glycoprotein (PGP) also referred
CC as multidrug transporter (MDR1) and nucleic acids encoding them.
CC The invention also includes fragments and biologically functional
CC variants of dog P-glycoprotein. PGP and their nucleic acids are
CC useful for determining the bioavailability of drugs and for
CC screening PGP inhibitors. They are useful for the diagnosis and
CC treatment of conditions characterised by PGP activity. By
CC reducing or increasing PGP activity in a cell, PGP nucleic acids
CC are used as oligonucleotide probes. Complements of PGP nucleic
CC acids are useful as antisense oligonucleotides, to induce a PGP
CC 'knockout' phenotype. They are used to prepare a non-human
CC transgenic animal, which are valuable as genetic models for
CC human diseases.
CC The present sequence is dog P-glycoprotein (PGP) cDNA. This
CC sequence is also referred as Genotype C cDNA. The
CC PGP enzyme functions as an efflux pump exporting small molecules
CC across the cell membrane. This enzyme is a member of the ABC
CC transporter family.
XX
SQ Sequence 4279 BP; 1294 A; 834 C; 1008 G; 1143 T; 0 other;

Alignment Scores:
Pred. No.: 1,69e-226 Length: 4279
Score: 2442.00 Matches: 469
Percent Similarity: 76.24% Conservative: 160
Best Local Similarity: 56.85% Mismatches: 178
Query Match: 59.87% Indels: 18
DB: Gaps: 4

US-09-873-409-2 (1-812) x AAD03488 (1-4279)
Qy 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
Db 1379 ATTGATGGACAGGACATTAGGACCATAAATGTAAGGCATCTTCGGGAAATTAATCGTGTG 1438
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
Db 1439 GTGAGTCAGGAGCCTGTGTGTTTCCACCGATAGCTGTAAGAACATTCGCTATGCGCGC 1498
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
Db 1499 GAAATGTCACCATGATGAGATTGAGAAAGCTGTTAAGAGACCAATGCCATATGATTTT 1558
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
```


Db 1559 ATCTGAAACTACCTAATAAATTTGACACTCTGGTGGAGAGAGGGGCCAGCTGAGT 1618
Qy 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
Db 1619 GTTGACAGAAACAGAGAAATGCCATTGCTCGGGCCCTGGTTCGCAACCCCAAGATTCTT 1678
Qy 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 121
Db 1679 CTGCTGGATGAGCAACGTCAGCTCTGGACACTGAAAGTGAAGCAGTGGTTCCAGGTGGCC 1738
Qy 122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
Db 1739 CTGGATAGGCCAGAAAGCCGCGACTACCATTTGTATGATAGCTCATCGTTTGTCTACAGTT 1798
Qy 142 ArgSerAlaAspLeuValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
Db 1799 CGTAATGCCGATGTCATTGCTGGTTTGTATGATGGATGTCATTGTGGAGAAAGAAATCAT 1858
Qy 162 AlaGluLeuMetAlaLysArgGlyLeuTyfTyfSerLeuValMetSerGlnAspLys 181
Db 1859 GATGAACTCATGAAGAGAGAGGCATTACTTCAAACTTGTCACAATGCAG---ACAAAGA 1915
Qy 182 LysAlaAspGluClnMetGluSerMetThrTyfSerThrGluArgLysThrAsnSerLeu 201
Db 1916 GGAAATGAATTCAGTTAGANAATGCCACTGGTGAATCCANAAGTCAAAAGTCATGCCTTG 1975
Qy 202 ProLeuHisSerValLysSerIleLysSerAspPheIle----- 214
Db 1976 GAAATG---TCTCCAAAAGATTCAGGTCAGTTTAAATAAAGAAAGATCAACTCGCAGG 2032
Qy 215 -----AspLysAlaGluLysThrGlnSerLysGluIle 226
Db 2033 AGTATACATGCACCACCAAGCCAGACAGAAAGCTTGGTACAAAAGAGGACTTGAATGAG 2092
Qy 227 SerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPhe 246
Db 2093 AATGTACCTCCAGTTTCCTTCCTCGAGGATTCGAAGCTCAACTCAACTGAATGGCCTTAT 2152
Qy 247 ValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIle 266
Db 2153 TTTGTGGTTGGTATATTTGTGCTATTATAAACGGAGGCTCGAACACCAITTTCAATA 2212
Qy 267 IlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAspLysThrThrLeuLysHis 285
Db 2213 ATATTTTCAAGGATTATAGGATCTTTACCCGAGATGAGGATCCTGAAACAAAACGACAG 2272
Qy 286 AspAlaGluIleTyfSerMetIlePheValIleLeuGlyValIleCysPheValSerTyf 305
Db 2273 AATAGTAACATGTTTCTGTATTGTTCTAGTCTTGGAAATATTCTTTTATTACATTT 2332
Qy 306 PheMetGlnGlyLeuPheTyfGlyArgAlaGlyGluIleLeuMetArgLeuArgHis 325
Db 2333 TTTCTCCAGGGCTTCACATTTGGCAAGCTGGGGAGATCCTCACTAAGCGGCTTCGATAC 2392
Qy 326 LeuAlaPheLysAlaMetLeuTyfGlnAspIleAlaTrpPheAspGluLysGluAsnSer 345
Db 2393 ATGGTTTTCAGATCCATGCTGACACAGATGTCAGCTGGTTTGATCACCCTAAAACACC 2452
Qy 346 ThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGly 365
Db 2453 ACTGGAGCATTCACACACAGGCTTGCCATGATGCGGCTCAAGTTAAAGGGGCTATAGT 2512
Qy 366 SerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSer 385
Db 2513 TCCAGCTTGGCTGCTATTACCCAGAAATATAGCAAAATCTTTGGGACAGGCATTATTATCC 2572
Qy 386 PheIleTyfGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaVal 405
Db 2573 TTAAATCATGGTTGGCAATTAACACTTTTACTTTAGCAATTTGTACCCATCATTCGAATA 2632
Qy 406 ThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeu 425
Db 425

Db 2633 GCAGGAGTTGTTGAAATGAAATGTTGTCTGGACACAGCACTGAAGAAGATAAGAAAGAGCTA 2692
Qy 426 LysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeu 445
Db 2693 GAAGGAGCTGGAGAGATTGCTACAGAACCCATCGAAAACTTCCGAACACTGTTGTTCTTTG 2752
Qy 446 ThrArgGluLysAlaPheGluGlnMetTyfGluGlnMetLeuGlnThrGlnHisArgAsn 465
Db 2753 ACTCGGAGCAGAAAGTTTGAATATACATGTATGCACAGAGTTTGCACAGTACCACATACA 2812
Qy 466 ThrSerLysLysAlaGlnIleIleGlySerCysTyfAlaPheSerHisAlaPheIleTyf 485
Db 2813 TCTTTGAGAAAGACACATCTTCGGGTCTCATTTCTATCACCAGCAGCAATGATGAT 2872
Qy 486 PheAlaTyfAlaAlaGlyPheArgPheGlyAlaTyfLeuIleGlnAlaGlyArgMetThr 505
Db 2873 TTTTCTATGCTGGCTGTTTCGGTTCCTACTTGGTGGCAATGAGTTTCATGAAC 2932
Qy 506 ProGluGlyMetPheIleValPheThrAlaIleAlaTyfGlyAlaMetAlaIleGlyLys 525
Db 2933 TTTTCAGGATGTTCTTTTGGTATCTCAGCTATTGCTTTGTTGGTCATGCGCAGTGGGCGAG 2992
Qy 526 ThrLeuValLeuAlaProGluTyfSerLysAlaLysSerGlyAlaAlaHisLeuPheAla 545
Db 2993 GTCAGTTCATTTGCTCTCGACTATGCCAAAGCCAAAGTATCAGCAGGCCACGTCATCATG 3052
Qy 546 LeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThr 565
Db 3053 ATCATTTGAAAGGCCCTCTGATTGACAGCTACAGCCCTCAGGCTCAAGCCAAATAGC 3112
Qy 566 CysGluGlyAsnLeuGluPheArgGluValSerPhePheTyfProCysArgProAspVal 585
Db 3113 TTGAAGCAAAATGTACATTTAAATGAGTCTGTTTCAACTATCCCACTCGCACACATC 3172
Qy 586 PheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGly 605
Db 3173 CCCGTCCTCAGGGGCTGAGCCTCAGGTGAAGAGGGCCAGCCTGCCCTCTGTAGGT 3232
Qy 606 SerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyfAspProVal 625
Db 3233 AGCAGTGGCTGTGGAGAGGACACAGTTGTTCTCAGCTCCTAGAGCGCTTCTATGACCCCTTG 3292
Qy 626 GlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArg 645
Db 3293 GCTGCTTCAGTGTCTAATTCATGGCAAGAGATAAAGACACCTGAATGTCCAGTGGCTCGA 3352
Qy 646 SerGlnIleAlaIleValProGlnClnProValLeuPheAsnCysSerIleAlaGluAsn 665
Db 3353 GCACACCTGGGGCATCGTGTCTCAGAGGCCCATCTCTGTTTGACTGCGAGCATTCGCCGAGA 3412
Qy 666 IleAlaTyfGlyAspAsnSerArgValProLeuAspGluIleLysGluAlaAlaAsn 685
Db 3413 ATTGCTATGGAGACAAACAGCGGCTCGTATCATCAAGAGATTATGCAGGACGCCAAG 3472
Qy 686 AlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyfAsnThrGlnValGly 705
Db 3473 GAGGCCAACATACACACTTCATCAGACACTCTCCCTGAGAAATACAAACACAGAGTAGA 3532
Qy 706 LeuLysGlyAlaGlnLeuSerGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeu 725
Db 3533 GACAAAGAACCCAGCTCTCTGTTGGCCAGAAACAGCCCATTTGCCATAGCTCGCCTCTT 3592
Qy 726 LeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSer 745
Db 3593 GTTAGAGCCTCATATTTTGTGTTGATGAAGTACATCAGCTCTCTGGATACAGAAAGT 3652
Qy 746 GluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValVal 765
Db 3653 GAAAGGTTTCCAGAGGCCCTGGCAAAAGCCAGAGAGGCGCGACCTGTCATTTGTATC 3712
Qy 766 ThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLys 785
Db 3713 GCCCACCCGCTTGTCCACCATCCAGAAATGCAGATTTAATAGTGTGTTTTCAGAAATGCCAAA 3772

QY 786 IleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeu 805
 Db 3773 GTCAAGGAGCATGGCACACATCAACAGCTGCTGGCCAGAAAGGCATCTATTTTCCATG 3832

QY 806 ValAsnAlaGlnSer 810
 Db 3833 GTCAGTGTCCAGGCT 3847

RESULT 6
 AAD03504
 ID AAD03504 standard; cDNA; 4279 BP.
 XX
 AC AAD03504;
 XX
 DT 13-JUN-2001 (first entry)
 XX
 DE Dog P-glycoprotein (PGP) allelic variant (Genotype A) cDNA.
 XX
 KW Dog; P-glycoprotein allelic variant; multidrug transporter; MDRI;
 KW drug bioavailability; transgenic animal; genetic model; ss.
 XX
 OS Canis familiaris.
 XX
 FH Key Location/Qualifiers
 FT CDS 17..3862
 FT /*tag= a
 FT /product= "Dog P-glycoprotein (PGP) allelic variant
 FT (Genotype A) protein"
 FT replace (607, C)
 FT /*tag= b

WO200123540-A2.
 PD 05-APR-2001.
 XX
 PF 28-SEP-2000; 2000WO-US26767.
 PR 28-SEP-1999; 99US-0156510.
 XX
 PA (GENT-) GENTEST CORP.
 XX
 PI Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;
 XX
 DR WPI; 2001-235373/24.
 DR P-PSDB; AAE00308.
 XX
 PT New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
 PT for determining the bioavailability of drugs and for screening for dog
 PT PGP inhibitors -
 XX
 PS Claim 9; Page 85-90; 11pp; English.
 XX
 CC The invention relates to dog P-glycoprotein (PGP) also referred
 CC as multidrug transporter (MDRI) and nucleic acids encoding them.
 CC The invention also includes fragments and biologically functional
 CC variants of dog P-glycoprotein. PGP and their nucleic acids are
 CC useful for determining the bioavailability of drugs and for
 CC screening PGP inhibitors. They are useful for the diagnosis and
 CC treatment of conditions characterised by PGP activity, by
 CC reducing or increasing PGP activity in a cell. PGP nucleic acids
 CC are used as oligonucleotide probes. Complements of PGP nucleic
 CC acids are useful as antisense oligonucleotides, to induce a PGP
 CC 'knockout' phenotype. They are used to prepare a non-human
 CC transgenic animal, which are valuable as genetic models for
 CC human diseases.
 CC The present sequence is dog P-glycoprotein (PGP) allelic variant
 CC (Genotype A) cDNA. The PGP enzyme functions as an efflux pump
 CC exporting small molecules across the cell membrane. This enzyme
 CC is a member of the ABC transporter family.
 XX
 SQ Sequence 4279 BP; 1295 A; 833 C; 1008 G; 1143 T; 0 other;

Alignment Scores:
 Pred. No.: 1.69e-226 Length: 4279
 Score: 2442.00 Matches: 469
 Percent Similarity: 76.24% Conservative: 160
 Best Local Similarity: 56.85% Mismatches: 178
 Query Match: 59.87% Indels: 18
 DB: 22 Gaps: 4

US-09-873-409-2 (1-812) x AAD03504 (1-4279)

QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
 Db 1379 ATTGATGACAGGACATTAGGACCAATAATGTAAGGCATCTTCGGGAAATTAATCTGGTGTG 1438

QY 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
 Db 1439 GTGAGTCAGGAGCCTGTGTGTTGCCACACGATAGCTGAAACATTCGCTATGCGCGC 1498

QY 42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
 Db 1499 GAAATATGTCACCATGATGAGATTGAGAAAGCTGTTAAGGAAGCAATGCCTATGATTTT 1558

QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyValaGlnMetSer 81
 Db 1559 ATCATGAAACTACCTAATAAATTTGACACTCTGGTTGGAGAGAGGCGCCAGCTGAGT 1618

QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
 Db 1619 GGTGACAGAAACAGAGAAATGCCATTGCTCGGCGCTGGTTGCGCAACCCCAAGATTCTT 1678

QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
 Db 1679 CTGCTGGATGAGGCAACGTCAGCTCTGGACACTGAAAGTGAAGCAGTGGTTCAGGTGCC 1738

QY 122 LeuGluLysAlaSerLysGlyArgThrIleValValAlaHisArgLeuSerThrIle 141
 Db 1739 CTGGATAAGGCCAGAAAGCGCGGACTACCAATTGTAGTCTCATCGTTGTCTACAGTT 1798

QY 142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
 Db 1799 CGTAATGCCGATGTCATTGCTGGTTTTCATGATGGAGTCATTGTGGAGAAAGGAATCAT 1858

QY 162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIleLys 181
 Db 1859 GATGAACATCATGAAAGAGAGAGGCATTTACTTCAAACTGTGCACAAATGCAG--ACAAGA 1915

QY 182 LysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeu 201
 Db 1916 GGAATGAAATTGAGTTAGAAATGCCACTGGTGAATCCAAAGTGAAAGTGAATGATCGCTTG 1975

QY 202 ProLeuHisSerValLysSerIleLysSerAspPheIle----- 214
 Db 1976 GAAATG---TCTCCAAAGATTTCAGGTCAGTTTAATAAAGAGAGATCAATCCGAGG 2032

QY 215 -----AspLysAlaGluGluSerThrGlnSerLysGluIle 226
 Db 2033 AGTATACATGCACCAAGGCCAAGACAGAAAGCTGTGTACAAAAGAGGACTTGAATGAG 2092

QY 227 SerLeuProGluValSerLeuLysIleLysLeuAsnLysProGluTyrProPhe 246
 Db 2093 AATGTACCTCCAGTTTCCTCTCGGAGGATTCGAAAGCTGAACCTCAACTGAATGCGCTTAT 2152

QY 247 ValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIle 266
 Db 2153 TTTGTGGTGTGTATATTTTGTGTATATAACCGAGGCGCTGCAACCAAGCATTTCAATA 2212

QY 267 IlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAspLysThrThrLeuLysHis 285
 Db 2213 ATATTTTCAAGGATATAGGATCTTTTACCAGAGATGAGGATCTCTGAAACCAAAACACGAG 2272

QY 286 AspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyr 305
 Db 2273 AATAGTAACATGTTTCTGTATGTTTCTAGTCTCTGGAATTAATTTCTTTTACATTT 2332

Qy 306 PheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluileLeuThrMetArgLeuArgHis 325
 Db 2333 TTCTCCAGGGCTTCCATATTTGGCAAGCTGGGAGATCTCTCACTAAGCGGCTTCGATAC 2392
 Qy 326 LeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTtrPheAspGluLysGluAsnSer 345
 Db 2393 ATGGTTTTTCAGATCCATGCTGAGACAGGATGTCACTGGTTGTAGACCTTAAACACACC 2452
 Qy 346 ThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGly 365
 Db 2453 ACTGAGCATTTGACACACAGGCTTGCCATGATGCGGCTCAAGTTAAGGGGCTATAGT 2512
 Qy 366 SerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSer 385
 Db 2513 TCCAGGCTTGCTGCTATTACCCAGAAATATAGCAATCTTGGGACAGGCATTTATATATCC 2572
 Qy 386 PheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaVal 405
 Db 2573 TTAATCTATGTTGGCAATTAACACTTTTACTCTTAGCAATTTGTACCCATCAATCAATA 2632
 Qy 406 ThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeu 425
 Db 2633 GCAGAGTTGTTGAAATGAAATGTTGCTGCAAGCACTCAAGATTAAGAAAGCTA 2692
 Qy 426 LysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeu 445
 Db 2693 GAAGAGCTGGGAGATTGCTACAGAAGCCATCGAAACTTCGAACTGTTGTTCTTTG 2752
 Qy 446 ThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsn 465
 Db 2753 ACTCGGAGCAGAAGTTTGAATACATGATGTACAGAGATTTGCAAGTACCATACAGAAAC 2812
 Qy 466 ThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyr 485
 Db 2813 TCTTTGAGGAAGACACATCTTCGGGGTCTCATTTCTATCACCCAGGCAATGATGAT 2872
 Qy 486 PheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThr 505
 Db 2873 TTTTCTATGCTGGCTGTTTCGGTTTGGTGTCTACTTGGTGCAAAATCAGTTTCATGAAC 2932
 Qy 506 ProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLys 525
 Db 2933 TTTTCAGGATGTTCTTTTGTGTTCTCAGCTATTGCTTTGGTGCCATGCGACGTGGGCGAC 2992
 Qy 526 ThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyValAlaHisLeuPheAla 545
 Db 2993 GTCAGTTCAATTTGCTCTGACTATGCAAGCCAAAGTATCAGCAGCCACGTCATCATG 3052
 Qy 546 LeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThr 565
 Db 3053 ATCATTTGAAAAGCCCTCTGATTGACAGTACAGCCCTCAGCGCTCAAGCCAAATACG 3112
 Qy 566 CysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspVal 585
 Db 3113 TTGGAAGGAATGTGACATTTAATGAGGTGCTGTTCAACTATATCCCACTCCAGCAGACATC 3172
 Qy 586 PheIleLeuArgGlySerLeuSerIleGluArgGlyLysThrValAlaPheValGly 605
 Db 3173 CCGGTCTCCAGGGGCTGAGCCTCGAGGTTGAGAGGGCCAGACGCTGGCCCTCGTAGGT 3232
 Qy 606 SerSerGlyCysGlyLysSerThrSerValGlnLeuGlnArgLeuTyrAspProVal 625
 Db 3233 AGCAGTGGCTGTGGGAAGACACAGTTGTTTACGTTCTTAGAGCGCTTCTATGACCCCTTG 3292
 Qy 626 GlnGlyGlnValLeuPheAspGlyValAlaPheAlaLysGluLeuAsnValGlnTyrLeuArg 645
 Db 3293 GCTGGTTCAGTGTAATGTATGTCGAAAGAGATAAAGCACCTGAATGTCAGTGGCTCCGA 3352
 Qy 646 SerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsn 665
 Db 3353 GCACACCTGGGCATGCTGTCTCAGGAGCCCATCTGTTTGACTGCAGCATTCGCCGAGAAC 3412

Qy 666 IleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsn 685
 Db 3413 ATTCCCTATGAGACAACACAGCCGGTCTGTATCATGAAGAGATTATCGAGCGACCAAG 3472
 Qy 686 AlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGly 705
 Db 3473 GAGGCCAACATACACCACTTCATCGAGACACTCCCTGAGAAATACCAACACAGAGTAGGA 3532
 Qy 706 LeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeu 725
 Db 3533 GACAAGGAACCCAGCTCTCTGGTGGCCAGAAACAGCGCATTTGCCATAGCTCGCGCTCTT 3592
 Qy 726 LeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSer 745
 Db 3593 GTTAGACAGCCTCATATTTTGTCTTTGGATGAAGCTACATCAGCTCTCGATACAGAAAGT 3652
 Qy 746 GluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValVal 765
 Db 3653 GAAAGGTTGTCCAGAGCCCTGGACAAAGCCAGAGAGCCGCGCACCTGCATTGTGATC 3712
 Qy 766 ThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLys 785
 Db 3713 GCCCACCCTTGTCCACCATCCAGATTCAGATTTAATAGTGTGTTTTCAGAAATGGCAA 3772
 Qy 786 IleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeu 805
 Db 3773 GTCAGGAGCATGGCACACATCAACAGCTGTGCGCCCAAGAGGCATCTATTTTTCATG 3832
 Qy 806 ValAsnAlaGlnSer 810
 Db 3833 GTCAGTGTCCAGGCT 3847
 RESULT 7
 ID AAD03505 standard; cDNA; 4279 BP.
 AC AAD03505;
 DT 13-JUN-2001 (first entry)
 DE Dog P-glycoprotein (PGP) allelic variant (Genotype B) cDNA.
 KW Dog; P-glycoprotein allelic variant; PGP; multidrug transporter;
 KW MDRI; drug bioavailability; transgenic animal; genetic model; ss.
 OS Canis familiaris.
 FH Key Location/Qualifiers
 FT CDS 17..3862
 FT /*tag= a
 FT /product= "Dog P-glycoprotein (PGP) allelic variant
 FT (Genotype B) protein"
 FT replace (91, T)
 FT /*tag= b
 FT replace (607, C)
 FT /*tag= c
 PN WO200123540-A2.
 XX
 PD 05-APR-2001.
 XX
 PF 28-SEP-2000; 2000WO-US26767.
 XX
 PR 28-SEP-1999; 99US-0156510.
 XX
 PA (GENT-) GENTEST CORP.
 XX
 PI Stocker PV, Steimel-crespi DT, Crespi CL, Reif TC, Patten CU;
 XX
 DR WPI; 2001-235373/24.
 DR P-PSDB; AAE00309.
 XX
 PT New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful

PT for determining the bioavailability of drugs and for screening for dog
 PT PGP inhibitors -

XX Claim 9; Page 93-99; 111pp; English.

CC The invention relates to dog P-glycoprotein (PGP) also referred
 CC as multidrug transporter (MDR1) and nucleic acids encoding them.
 CC The invention also includes fragments and biologically functional
 CC variants of dog P-glycoprotein. PGP and their nucleic acids are
 CC useful for determining the bioavailability of drugs and for
 CC screening PGP inhibitors. They are useful for the diagnosis and
 CC treatment of conditions characterised by PGP activity, by
 CC reducing or increasing PGP activity in a cell. PGP nucleic acids
 CC are used as oligonucleotide probes. Complements of PGP nucleic
 CC acids are useful as antisense oligonucleotides, to induce a PGP
 CC 'knockout' phenotype. They are used to prepare a non-human
 CC transgenic animal, which are valuable as genetic models for
 CC human diseases.

CC The present sequence is dog P-glycoprotein (PGP) allelic variant
 CC (Genotype B) cDNA. The PGP enzyme functions as an efflux pump
 CC exporting small molecules across the cell membrane. This enzyme
 CC is a member of the ABC transporter family.

XX SQ Sequence 4279 BP; 1296 A; 833 C; 1008 G; 1142 T; 0 other;

Alignment Scores:

Pred. No.:	1,69e-226	Length:	4279
Score:	2442.00	Matches:	469
Percent Similarity:	76.24%	Conservative:	160
Best Local Similarity:	56.85%	Mismatches:	178
Query Match:	59.87%	Indels:	18
DB:	22	Gaps:	4

US-09-873-409-2 (1-812) x AAD03505 (1-4279)

QY	2	ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal	21
DB	1379	ATTGATGGACAGACATTAGGACCAATAAATGTAAAGCATCTTCGGGAAATTAATCGTG	1438
QY	22	ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg	41
DB	1439	GTGAGTCAGGAGCTGTGTGTTGGCCACCATAGTAGTGAAACATTCGCTATGCGCCG	1498
QY	42	AspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe	61
DB	1499	GAATAATGTCACCATGATGAGATTGAGAAAGCTGTTAAGGAGCCCAATGCCATGATGTTT	1558
QY	62	IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer	81
DB	1559	ATCATGAAACTACCTAATAAATTGACACTCTGTGTTGGAGAGAGAGGGGCCAGCTGAGT	1618
QY	82	GlyGlyGlnLysGlnArgIleAlaIleAlaAatGAlaLeuValArgAsnProLysIleLeu	101
DB	1619	GTGTGGACAGAAACAGAAATCGCCATTCGTGGGCCCTGTGTCGCAACCCCAAGATCTT	1678
QY	102	IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla	121
DB	1679	CTGCTGGATGAGGCAACGTCAGCTCTGGACACTGAAAGTGAACAGCTGTTTCAGGTGGCC	1738
QY	122	LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle	141
DB	1739	CTGGATTAAGCCAGAAAAGCCGAGCTACCATTTGTGATAGCTCATCTGTTGTCTACAGTT	1798
QY	142	ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis	161
DB	1799	CGTAATGCCGATGTCATTCGCTGTTTGTATGATGGAGTCAATGTTGGAGAAAGAAATCAT	1858
QY	162	AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIleLys	181
DB	1859	GATGAACATCATGAAGAGAGAGGCAATTAATCTCAAACTTGTCAATATGCAG--ACAAGA	1915
QY	182	LysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeu	201
DB		::: ::::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: :::	

DB	1916	GGAAATGAAATTTAGTTAGATAAATGCCACTGGTGAATCCAAAAGTGAAGATGATGCCTTG	1975
QY	202	ProLeuHisSerValLysSerIleLysSerAspPheIle-----	214
DB	1976	GAAATG--TCTCCAAAAGATTTCAGGGTCCAGTTTAAATAAAAAGAGATCAACTCGCAGG	2032
QY	215	-----AspLysAlaGluGluSerThrGlnSerLysGluIle	226
DB	2033	AGTATACATGCACCAAGGCGCAAGAAAGCTTGGTACAAAAGAGAGACTTGAATGAG	2092
QY	227	SerLeuProGluValSerLeuLysIleLeuLysLeuAsnLysProGluTrpProPhe	246
DB	2093	AATGACCTCCAGTTCTCTTCGGAGGATTTGAAGCTGAACCTCAACTGAATGGCCTTAT	2152
QY	247	ValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIle	266
DB	2153	TTTGTGGTTGGTATATTTTGTCTATTATAAACGGAGGCTGCAACACGACCATTTTCAATA	2212
QY	267	IlePheAlaLysIleIleThrMetPhe--GlyAsnAsnAspLysThrThrLeuLysHis	285
DB	2213	ATATTTTCAAGGATTTATAGGATCTTTACCCGAGATGAGGATCTCTGAAACAAACAGCAG	2272
QY	286	AspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyr	305
DB	2273	AATAGTAACATGTTTCTGTATTTGTTCTTGTGCTTGAATTTATTTTATTACATTT	2332
QY	306	PheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHis	325
DB	2333	TTCTCCAGGGCTTCACATTTGGCAAGCTGGGAGATCTCTCACTAAGGGCTTCGATAC	2392
QY	326	LeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSer	345
DB	2393	ATGGTTTTTCAGATCCATGCTGAGACAGGATGTGAGCTGTTGTAGTGCCTTAAACACACC	2452
QY	346	ThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyValaThrGly	365
DB	2453	ACTGAGCATTTGACAAACAGGCTTCCCAATGATGGCTCAAGTTTAAAGGGCTTAGGT	2512
QY	366	SerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSer	385
DB	2513	TCCAGCTTGCTGTCTATTTACCCAGAAATATAGCAATCTTGGACAGGCTATATATACC	2572
QY	386	PheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaVal	405
DB	2573	TTAATCTATGTTGGCAATTAACACTTTTACTTTAGCAATTTGTACCCTCATITCAATA	2632
QY	406	ThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeu	425
DB	2633	CGAGAGTTGTTGAAATGAAATGTTGTTGGACAAAGCACTGAAAGATAAGAGAGCTTA	2692
QY	426	LysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeu	445
DB	2693	GAAGAGCTGGGAAGATTGCTACAGAGCCCATCGAAAACCTCCGAACTGTGTTCTTTTG	2752
QY	446	ThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsn	465
DB	2753	ACTCGGAGCAGAAAGTTTGAATACATGATGACAGAGTTTGCAAGTACCACATACAGAAAC	2812
QY	466	ThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyr	485
DB	2813	TCCTTTGGAAGAGCACATCTTCGGGTGTCATTTTCTTCCAGCCGCAATGATGAT	2872
QY	486	PheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThr	505
DB	2873	TTTTCTATGCTGGCTGTTTCGGTTTGGTGGCTACTTGGTGGCAAAATGAGTTCATGAAC	2932
QY	506	ProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLys	525
DB	2933	TTTCAGGATGTTCTTTGTGATTTCTCAGCTATTTGTTGTCCTTGGCAGTGGCGGAG	2992
QY	526	ThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAla	545
DB	2993	GTGATTTGCTTCTGACTATGCTTATGCCAAAGTATCAGCAGCCCACTGTCATG	3052


```
Qy 546 LeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThr 565
    ::::::::::|||
Db 3053 ATCATTTAAAAAGCCCTCTGATTGACAGCTACAGCCCTCAGCGCTCAAGCAAAATACG 3112

Qy 566 CysGluGlyAsnLeuGluPheArgGluValSerPheThrProCysArgProAspVal 585
    ::::::::::|||
Db 3113 TTGGAAGGAATGTGACATTTTAATGAGGTGCTGTTCAACTATCCACATCGACACATC 3172

Qy 586 PheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGly 605
    ::::::::::|||
Db 3173 CCCGTGCTCCAGGGCTGAGCTTCGAGGTGAAGAGGCCACACCTGGCCCTCTAGGT 3232

Qy 606 SerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuThrAspProVal 625
    ::::::::::|||
Db 3233 AGCAGTGGCTGTGGGAAGACAGCATGTTTTCAGCTCCTAGAGCGCTTCTATGACCCCTTG 3292

Qy 626 GlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArg 645
    ::::::::::|||
Db 3293 GCTGGTTCAGTCTTAATTGATGCGCAAGAGATATAAGCACCTGAATGTCCAGTGGCTCCGA 3352

Qy 646 SerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsn 665
    ::::::::::|||
Db 3353 GCACCTGGGGCATGCTGCTCAGAGAGCCCATCTCTTTGACTGGAGCATTCGCCAGAAC 3412

Qy 666 IleAlaIleThrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaIleAsn 685
    ::::::::::|||
Db 3413 ATTGCTATGGAGACACACAGCCGGTCTGATCATCAAGAGATTATGCAGGCAGCCAAAG 3472

Qy 686 AlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysThrAsnThrGlnValGly 705
    ::::::::::|||
Db 3473 GAGGCGCAACATACACACTTTCATCAGACACTCCCTCGAAGAAATACAAACACAGATAGGA 3532

Qy 706 LeuLysGlyAlaGlnLeuSerGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeu 725
    ::::::::::|||
Db 3533 GACAAGGAACCCAGCTCTCTGGTGGCCGAAGAACGCGCATTTGCCATGCTCGCGCTCT 3592

Qy 726 LeuGlnLysProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSer 745
    ::::::::::|||
Db 3593 GTTAGACAGCCCTCATATTTTCTTTGGATGAAGCTACATCAGCTCTGGATACAGAAAGT 3652

Qy 746 GluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValVal 765
    ::::::::::|||
Db 3653 GAAAAGTTGTCCAAGAACCCCTGGCAAAAGCCAGAGAGCGCCACCTGTCATTTGTGATC 3712

Qy 766 ThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLys 785
    ::::::::::|||
Db 3713 GCCCACCCTTGTCCACCATCCAGAAATGAGATTAAATAGTGTGTTTCAGATGGCAAA 3772

Qy 786 IleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleThrPheLysLeu 805
    ::::::::::|||
Db 3773 GTCAAGGAGCATGGCACATCAACAGCTGCTGGCCAGAAAGGCATCTATTTTCCATG 3832

Qy 806 ValAsnAlaGlnSer 810
    ::::::::::|||
Db 3833 GTCAAGTGTCCAGCT 3847

RESULT 8
AAF86127
ID AAF86127 standard; cDNA; 4186 BP.
XX
AC AAF86127;
XX
DT 25-JUN-2001 (first entry)
XX
DE Cynomologous monkey P-glycoprotein cDNA.
XX
KW Cynomologous monkey; P-glycoprotein; PGP; multidrug transporter; MDR1;
KW efflux pump; ss.
XX
OS Macaca fascicularis.
XX
FH Key Location/Qualifiers
```

```
CDS 100..3942
FT /*tag= a
FT /product= "PGP"
FT /note= "P-glycoprotein"
XX
PN WO200123565-A1.
XX
PD 05-APR-2001.
XX
PF 28-SEP-2000; 2000WO-US26592.
XX
PR 28-SEP-1999; 99US-0156921.
PR 12-OCT-1999; 99US-0158818.
XX
PA (GENT-) GENTEST CORP.
XX
PI Stocker PJ, Steimel-Crespi DT, Creepi CL;
XX
DR WPI; 2001-316136/33.
DR P-PSDB; AAB81064.
XX
Novel isolated nucleic acid encoding cynomologous monkey P-glycoprotein
(PGP) and homologous PGP polypeptides are useful for predicting
bioavailability of compound and increasing PGP transporter activity in
cell
Example 1; Page 51-57; 84pp; English.
This invention relates to a polynucleotide sequence encoding a
cynomologous monkey P-glycoprotein (PGP), and an allelic variant of the
PGP protein. PGP, also known as multidrug transporter, MDR1 is a member
of the ABC transporter superfamily. The enzyme serves as an efflux pump
exporting small molecules across the cell membrane. The invention
includes a cynomologous monkey (Macaca fascicularis) PGP coding sequence
and protein, and also that of an allelic variant. The PGP polynucleotide
sequence is useful for increasing PGP transporter activity in a cell.
Antisense sequences of the cDNA are useful for inhibiting PGP transport
activity in a mammalian cell. They may also be used for increasing the
bioavailability of a drug. The present sequence represents cDNA encoding
cynomologous monkey P-glycoprotein.
SQ Sequence 4186 BP; 1226 A; 799 C; 1039 G; 1122 T; 0 Other;
```

```
Alignment Scores:
Pred. No.: 2,04e-226 Length: 4186
Score: 2441.00 Matches: 465
Percent Similarity: 76.73% Conservative: 168
Best Local Similarity: 56.36% Mismatches: 174
Query Match: 59.84% Indels: 18
DB: 22 Gaps: 5
```

```
US-09-873-409-2 (1-812) x AAF86127 (1-4186)
Qy 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
    |||||
Db 1456 GTTGATGACAGCATATTAGGACCATAAACGTAAAGTTTCTACGGGAAATCATCGGTGTG 1515

Qy 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
    |||||
Db 1516 GTGAGTCAGGAACCTGTATTGTTGCCACCATGATAGTGAAGAAATTCGCTATGTCGT 1575

Qy 42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
    |||||
Db 1576 GAAGATGTCACCATGATGAGATTGAGAAAGCTGTCAAGAGCAATGCTATGACTTT 1635

Qy 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
    |||||
Db 1636 ATCATGAAACTCCCTCAGAAATTTGACACCTCTGTTGGAGAGAGAGGGCCCGCTGAGT 1695

Qy 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
    |||||
Db 1696 GGTGGGAGAGACAGAGGATGCCATTGCAGTGCCTTGGTTCGCAACCCCAAGATCCTC 1755
```


Qy	102	Ileu	Asp	Glu	Ala	Thr	Ser	Ala	Leu	Asp	Ser	Glu	Ser	Leu	Ser	Ala	Val	Gln	Ala	Ala	121
Db	1756	CTG	TGG	AC	GAG	GCC	ACG	T	CAG	CTT	TGG	CA	CAC	AA	AGT	TGA	AGC	AGG	CTG	TGG	1815
Qy	122	Leu	Glu	Leu	Ala	Ser	Leu	Gly	Arg	Thr	Thr	Ile	Val	Ala	Ile	Asn	Leu	Ser	Thr	Ile	141
Db	1816	CTG	AT	AG	CC	CA	AAA	AGT	CG	ACC	CA	ATT	G	TG	AT	AGC	T	C	T	CG	1875
Qy	142	Arg	Ser	Ala	Asp	Leu	Ile	Val	Thr	Leu	Leu	Ser	Asp	Gly	Met	Leu	Ala	Glu	Leu	Gly	161
Db	1876	CGT	AA	T	CCG	AC	GT	C	T	CG	T	GGT	T	CG	AT	G	T	GC	GA	AAA	1935
Qy	162	Ala	Glu	Leu	Met	Ala	Leu	Arg	Gly	Leu	Tyr	Tyr	Ser	Leu	Val	Met	Ser	Gln	Asp	Ile	180
Db	1936	GAT	G	AGC	T	CA	T	GA	AG	AAA	G	CA	CTT	T	ACT	T	CA	AA	CT	T	1995
Qy	181	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	191
Db	1996	AAT	G	AA	T	T	CA	AAA	AT	GC	AGC	T	GA	T	GA	T	CC	AAA	AGT	GA	2055
Qy	192	Tyr	Ser	Thr	Glu	Arg	Leu	Ser	Thr	Asn	Ser	Leu	-----	-----	-----	-----	-----	-----	-----	-----	206
Db	2056	ATG	T	CTT	CA	CAT	G	AT	T	C	CA	G	T	CT	AA	T	A	GA	AAA	AG	2115
Qy	207	Lys	Ser	Ile	Lys	Ser	Asp	Phe	Ile	Asp	Leu	Ala	Glu	Ser	Thr	Gln	Ser	Lys	Glu	Ile	226
Db	2116	CGT	G	AT	CA	CA	AG	CC	AA	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	2169
Qy	227	Ser	Leu	Pro	Glu	Val	Ser	Leu	Leu	Ser	Ile	Leu	Ser	Leu	Asn	Lys	Pro	Glu	Tyr	Pro	246
Db	2170	AGT	A	T	AC	CT	CC	CA	G	T	T	T	CG	AG	GA	T	T	A	GA	CT	2229
Qy	247	Val	Val	Leu	Gly	Thr	Leu	Ala	Ser	Val	Leu	Asn	Gly	Thr	Val	Ile	Pro	Val	Phe	Ser	266
Db	2230	TTT	G	T	T	G	T	G	T	AT	T	T	T	G	C	CA	T	AT	A	A	2289
Qy	267	Ile	Phe	Ala	Lys	Ile	Ile	Thr	Met	Phe	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	285
Db	2290	AT	A	T	T	T	CA	A	AG	A	T	T	A	G	A	A	T	G	A	T	2349
Qy	286	Asp	Ala	Glu	Ile	Tyr	Ser	Met	Ile	Phe	Val	Ile	Leu	Gly	Val	Ile	Cys	Phe	Val	Ser	305
Db	2350	AAT	AGT	AA	CT	T	G	T	T	T	CT	AT	T	T	T	CT	AG	T	T	T	2409
Qy	306	Phe	Met	Gln	Gly	Leu	Phe	Tyr	Gly	Arg	Ala	Gly	Glu	Ile	Leu	Thr	Met	Arg	Leu	Arg	325
Db	2410	TT	C	CT	C	AGG	CT	T	CA	T	T	G	C	A	A	G	T	C	C	A	2469
Qy	326	Leu	Ala	Phe	Lys	Ala	Met	Leu	Tyr	Gln	Asp	Ile	Ala	Tyr	Phe	Asp	Glu	Gly	Leu	Asn	345
Db	2470	ATG	G	T	T	T	TC	G	AT	CT	C	CA	G	AG	AT	G	T	G	AG	CT	2529
Qy	346	Thr	Gly	Leu	Thr	Thr	Ile	Leu	Ala	Ile	Asp	Ile	Ala	Gln	Ile	Gln	Gly	Ala	Thr	Gly	365
Db	2530	ACT	G	G	CA	T	T	GA	CT	CG	CA	A	T	GA	T	GT	CT	CA	AGT	T	2589
Qy	366	Ser	Arg	Ile	Gly	Val	Leu	Thr	Gln	Asn	Ala	Thr									

Db	2830	ACTCAGGACGAGAAGTTTGAACATATGTATGATCATCAGAGTTTGCAGGTACCATTCACAGAAC	2889
Qy	466	ThrSerLysLysAlaGlnIlelleGlySerCysTyrAlaPheSerHisAlaPheIleTyr	485
Db	2890	TCATTGAGGAAGCACACATCTTTGGATCATCGTTTCTTCACGCGAGCAATGATGTAT	2949
Qy	486	PheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThr	505
Db	2950	TTTTCCTATGCTGGATGTTTCCGGTTTGAGCGCTACTTGGTGGCACATAGTCTCATGAGC	3009
Qy	506	ProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLys	525
Db	3010	TTTGAGGATGTTCTGTGTAGTATTTTCAGCTGTGTCTTTGTGGTCATGCCGTGGGGCAA	3069
Qy	526	ThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAla	545
Db	3070	GTCAAGTTCATTTGCTCTCACTATGCCAAGCCAAAGTATCAGCAGCCACATCATCATG	3129
Qy	546	LeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThr	565
Db	3130	ATCATTTGAAAAACCCCTTTTGATTGACAGTACAGCACAGAAGGCCCTAAAGCCGAACACA	3189
Qy	566	CysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspVal	585
Db	3190	TTGGAAGGAATGTCACTTTAATGAAGTTGTATTCAACTATCCACCOCGACTGGACATC	3249
Qy	586	PheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGly	605
Db	3250	CCAGTGTCTTCAGGGCTGAGCCTTGAAGTGAAGAAGGCCAGACGCTGGCCCTGGTGGC	3309
Qy	606	SerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProVal	625
Db	3310	AGCAGTGGCTGTGGGAAGACGACGGTGTCTCCAGCTCTCTGGAGCGGTCTATGACCCCTTG	3369
Qy	626	GlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTyrLeuArg	645
Db	3370	CGCGGGAAGTGTCTTTCAGCGCAAGAAATAAAGCAACTGAATGTTTCAGTGGCTCCGA	3429
Qy	646	SerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsn	665
Db	3430	GCACACTGGGCATCGTGTCCAGAGGCCATCTCTGTGTGACTCGAGCATTTAGTGAGAAC	3489
Qy	666	IleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsn	685
Db	3490	ATTGCCCTATGGACACAACAGCCGGTGTGTTCACAGGAAGAGATCGTGAGGCGACCAAG	3549
Qy	686	AlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGly	705
Db	3550	GAGGCCAATATACACGCCTTTCAGAGTCACTGCGCTTAATAATATAGCACACAGTAGTAGA	3609
Qy	706	LeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeu	725
Db	3610	GACAAAGGAAGTCACTGCTCTCTGGTGCCAGAAACCAACGATTCGCATAGCTCTGTGGCCCTT	3669
Qy	726	LeuGlnLysProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSer	745
Db	3670	GTTAGACAGCCTCATATTTTGTCTTTTGGNTGAGGCCACATCAGCTCTGGATACAGAAAGT	3729
Qy	746	GluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValVal	765
Db	3730	GAAAAAGTTGTCCAAAGACCCCTGGACAAAGCCAGAGAAGGCCGTACCTGCATTTGTGATT	3789
Qy	766	ThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLys	785
Db	3790	GCTCACCGCCTGTCCACCATCCAGAAATGCAGACTTAATAGTGGTGTTCACAAATGGCAGA	3849
Qy	786	IleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeu	805
Db	3850	GTCAAGAGACGCGCACACATCAGCAGCTGTCTGGCACAGAAGGCCATCTATTTTTCATGT	3909
Qy	806	ValAsnAlaGlnSer	810

Dbb 2479 ATGGTTTCCTGATCCATGCTCAGACAGGATGTGAGCTGTTGATGACCCCTAAACACACC 2538
Qy 346 ThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGly 365
Dbb 2539 ACTGGAGCATTCGACTACCAAGGCTCGCAATGATGCTGCTCAAGTTAAAGGGGCTATAGT 2598
Qy 366 SerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSer 385
Dbb 2599 TCCAGGCTTGCTATATATACCCAGATATAGCAAAATCTTGGGACAGGAATAATATATCC 2658
Qy 386 PheIleThrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaVal 405
Dbb 2659 TTAATCTATGTTGGCACTGACCTGTACTCTTAGCAATGTACCCATCATGTGCAATA 2718
Qy 406 ThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeu 425
Dbb 2719 GCAGGAGTTGTTGAATGAATGTTGTTGTCGACAGCACTGAAAGATGAAGAAGACTA 2778
Qy 426 LysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeu 445
Dbb 2779 GAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACCTTCGAACTGTGTTCTTTG 2838
Qy 446 ThrArgGluLysAlaPheGluGlnMetThrGluMetLeuGlnThrGlnHisArgAsn 465
Dbb 2839 ACTCAGAGCAGCAAGTTTGAACATATGATGATGATGACAGTTTGCAGGTACCATCAGAAC 2898
Qy 466 ThrSerLysLysAlaGlnIleIleGlySerCysThrAlaPheSerHisAlaPheIleTyr 485
Dbb 2899 TCTTTGAGGAAGCACACATCTTTGGAATCAGCTTTCTTCAGCGAGGCAATGATGAT 2958
Qy 486 PheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrIleuIleGlnAlaGlyArgMetThr 505
Dbb 2959 TTTTCTATGCTGGATGTTTCCGTTTGGAGCCTACTTGTGTGCACATGATGCTCATGAGC 3018
Qy 506 ProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyValaMetAlaIleGlyLys 525
Dbb 3019 TTTGAGATGCTCTGTAGTATTTTTCAGCTGTGTCTTTGGTCCATGGCCGTGGGGCAA 3078
Qy 526 ThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyValaAlaHisLeuPheAla 545
Dbb 3079 GTCAAGTTTCCTCTGCTGACTATGCCAAAGCCAAAGTATCAGCAGCCCATCATCATG 3138
Qy 546 LeuLeuGluLysLysProAlaIleAspSerArgSerGlnGluGlyLysLysProAspThr 565
Dbb 3139 ATCATTTGAAAAACCCCTTTGATTGACGTACAGCACAGCAAGGCCCTAAAGCCGACACA 3198
Qy 566 CysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspVal 585
Dbb 3199 TTGAAGGAATATGTCATTTAATGAAGTTGTATTCACTATCCACCCGACTGGACATC 3258
Qy 586 PheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGly 605
Dbb 3259 CNAAGTCTTCAGGGGCTGAGCTGGAAAGTGAAGAGGCCAGACGCTGGCCCTGGTGGGC 3318
Qy 606 SerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProVal 625
Dbb 3319 AGCAGTGGCTGTGGGAAGACGACGGTGGTCCAGCTCTGGAGCGGTCTATGACCCCTTG 3378
Qy 626 GlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArg 645
Dbb 3379 GCGGGAAAGTGTGCTGTACGGCAAAAGAAATAAGAACTGTAATGTTAGTGGCTCCGA 3438
Qy 646 SerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsn 665
Dbb 3439 GCACACCTGGGCATCGTGTCCAGGAGCCCATCTCTGTTGACTGCAGCATTAGTGAGAAC 3498
Qy 666 IleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsn 685
Dbb 3499 ATTGCCATATGGACACACAGCCGGTGTGTACAGGAAGAGATCGTGGAGGGCAGCAAG 3558
Qy 686 AlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGly 705
Dbb 3559 GAGGCCAATATACACGCCCTTCATCGAGTCACTGCCCTAATAATATAGTACACAGTAGGA 3618

Qy 706 LeuLysGlyAlaGlnLeuSerGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeu 725
Dbb 3619 GACAAAGGAACCTCAGCTCTCTGGTGGCCAGAAACACGATTCGCATAGCTGTCGCCCTT 3678
Qy 726 LeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSer 745
Dbb 3679 GTTACAGACGCTCATATTTGCTTTTGGATGAAGCCACATCAGCTCTGGATACAGAAAGT 3738
Qy 746 GluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValVal 765
Dbb 3739 GAAAGGTGTTCCAGAAAGCCCTGGACAAAGCCAGAAAGCCGATACCTGCTGCTGATT 3798
Qy 766 ThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLys 785
Dbb 3799 GCTCACCCTGCTCCACCATCCAGAAATGACAGCTTAAATAGTGTGTTTCAGAAATGCCACA 3858
Qy 786 IleLysGluGlnGlyThrHisGlnLeuLeuArgAsnArgAspIleTyrPheLysLeu 805
Dbb 3859 GTCAAGGAGCAGCGCACACATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTTCAATG 3918
Qy 806 ValAsnAlaGlnSer 810
Dbb 3919 GTCACTGTCAGGCT 3933
RESULT 10
AAD03506
ID AAD03506 standard; cDNA; 4279 BP.
XX AAD03506;
AC AAD03506;
XX 13-JUN-2001 (first entry)
XX Dog P-glycoprotein (PGP) allelic variant (Genotype D) cDNA.
XX Dog; P-glycoprotein allelic variant; PGP; multidrug transporter;
KW MDRI; drug bioavailability; transgenic animal; genetic model; ss.
XX Canis familiaris.
PH Key Location/Qualifiers
FT CDS 17..3862
FT /*tag= a
FT /product= "Dog P-glycoprotein (PGP) allelic variant
FT (Genotype D) protein"
FT allele replace (91, T)
FT /*tag= b
FT allele replace (607, C)
FT /*tag= c
FT allele replace (1001, T)
FT /*tag= c
FT allele replace (3458, A)
FT /*tag= c
XX WO200123540-A2.
PN 05-APR-2001.
XX 28-SEP-2000; 2000WO-US26767.
XX 28-SEP-1999; 99US-0156510.
XX (GENT-) GENTEST CORP.
PA Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;
PI WPI; 2001-235373/24.
XX P-PSDB; AAE00310.
PT New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
PT for determining the bioavailability of drugs and for screening for dog
PT PGP inhibitors -
XX

PS Claim 9; Page 102-107; 111pp; English.

CC The invention relates to dog P-glycoprotein (PGP) also referred
CC as multidrug transporter (MDR1) and nucleic acids encoding them.
CC The invention also includes fragments and biologically functional
CC variants of dog P-glycoprotein. PGP and their nucleic acids are
CC useful for determining the bioavailability of drugs and for
CC screening PGP inhibitors. They are useful for the diagnosis and
CC treatment of conditions characterised by PGP activity, by
CC reducing or increasing PGP activity in a cell. PGP nucleic acids
CC are used as oligonucleotide probes. Complements of PGP nucleic
CC acids are useful as antisense oligonucleotides, to induce a PGP
CC 'knockout' phenotype. They are used to prepare a non-human
CC transgenic animal, which are valuable as genetic models for
CC human diseases.
CC The present sequence is dog P-glycoprotein (PGP) allelic variant
CC (Genotype D) cDNA. The PGP enzyme functions as an efflux pump
CC exporting small molecules across the cell membrane. This enzyme
CC is a member of the ABC transporter family.

XX SQ Sequence 4279 BP; 1296 A; 833 C; 1009 G; 1141 T; 0 other;

Alignment Scores:

Pred. No.:	2,11e-226	Length:	4279
Score:	2441.00	Matches:	469
Percent Similarity:	76.24%	Conservative:	160
Best Local Similarity:	56.85%	Mismatches:	178
Query Match:	59.84%	Indels:	18
DB:	22	Gaps:	4

US-09-873-409-2 (1-812) x AAD03506 (1-4279)

Qy	2	ValAspGluAsnAspIleArgAlaLeuAenValArgHisTyrArgAspHisIleGlyVal	21
Db	1379	ATTGATGACAGGACATTAGGACCAATAATGTAAAGCACTTCGCGGAATTAAGTGGTG	1438
Qy	22	ValSerGlnGluProValLeuPheGlyThrIleSerAsnAsnIleLysTyrGlyArg	41
Db	1439	GTGAGTCAGGAGCCCTGTGTTGTCACACGATAGCTGAAACATTCGCTATGCGCCG	1498
Qy	42	AspAspValThrAspGluMetGluArgAlaAlaAArgGluAlaAsnAlaTyrAspPhe	61
Db	1499	GAATAATGTCACCATGGATGAGATGAGAAAGCTGTAAAGAAAGCCCAATGCCCTATGATTT	1558
Qy	62	IleMetGluPheProAsnLysPheAenThrLeuValGlyGlyAlaGlnMetSer	81
Db	1559	ATCATGAACACTACCTAATAATTTTGACCTCTGGTTGGAGAGAGGGCCCGCTGAGT	1618
Qy	82	GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu	101
Db	1619	GCTGGACAGAAACACAGAGATGCCATTGCTCGGGCCCTGTTCCGAACCCCAAGATTCIT	1678
Qy	102	IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla	121
Db	1679	CTGCTGGATGAGCAACGTCAGCTCTGGACACTGGAAGTGAAGCAGCATGTTTCAGGTGGCC	1738
Qy	122	LeuGluLysAlaSerLysGlyArgThrIleValValAlaHisArgLeuSerThrIle	141
Db	1739	CTGGATAGCGCCAGAAAGCGGACTACATTTGTATGATGCTCATCTGTTGTCACGTT	1798
Qy	142	ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis	161
Db	1799	CGTAATGCCGATGCTATTGCTGTTTGTATGATGGAGTCATGTTGGAGAAAGAAATCAT	1858
Qy	162	AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIleLys	181
Db	1859	GATGAACCTCATGAAGAGAGAGGCGCATTTACTTCAAACTTGTGCACATGTCAG---ACAAGA	1915
Qy	182	LysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeu	201
Db	1916	GGAAATGAATGAGTTGAAATGCCATCTGGTGAATCCAAAGTCAAAAGTGAATGCTG	1975
Qy	202	ProLeuHisSerValLysSerIleLysSerAspPheIle-----	214

Db	1976	GAATG---TCTCCAAAAGATTTCAGGGTCCAGTTTATAATAAAAGAGATCAACTCGCAGG	2032
Qy	215	-----AspLysAlaGluGluSerThrGlnSerLysGluIle	226
Db	2033	AGTATACATGCACCACCAAGGCCACAGACAGAAAGCTTGGTACAAAAGAGGACTTGAATGAG	2092
Qy	227	SerLeuProGluValSerLeuLeuLysIleLeuLysLeuAenLysProGluTyrProPhe	246
Db	2093	AATGTACCTCCAGTTTCTCTCGGAGGATCTGAAGCTGAACCTCAACTCAATGGCGCTTAT	2152
Qy	247	ValValLeuGlyThrLeuAlaSerValLeuAenGlyThrValHisProValPheSerIle	266
Db	2153	TTTGTGGTGTGTATATTTGTGCTATTATAAACGGAGCGCTGCAACACGATTTTCAATA	2212
Qy	267	IlePheAlaLysIleIleThrMetPhe---GlyAenAsnAspLysThrThrLeuLysHis	285
Db	2213	ATATTTTCAAGATATATAGGATCTTTACCCGAGATGAGGATCTCTGAACCAAAACACAG	2272
Qy	286	AspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyr	305
Db	2273	AATAGTAACATGTTTCTGTATTTCTAGTCCTTGAATTAATTTCTTTTATTACATTT	2332
Qy	306	PheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHis	325
Db	2333	TTCTCTCCAGGGCTTCACATTTTGGCAAGCTGGGGAGATCTCTACCTAAAGCGGCTTCGATAC	2392
Qy	336	LeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGluLysGluAsnSer	345
Db	2393	ATGGTTTTCAGATCCATGCTGAGACAGGATGTACGCTGGTTGTATGATGCCCTTAAACACC	2452
Qy	346	ThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGly	365
Db	2453	ACTGGAGCATTTGACAAACCCAGGCTTGCCCAATGATGGCGCTCAAGTTAAAGGGGCTATAGT	2512
Qy	366	SerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSer	385
Db	2513	TCCAGGCTTGCTGTCTATTACCAGAAATATAGCAATCTTGGGACAGGCATTAATATATCC	2572
Qy	386	PheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaVal	405
Db	2573	TTAATCTATGTTGGCAATTAACATTTTCTCTTAGCAATTTGTACCCATCATTCGATTA	2632
Qy	406	ThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeu	425
Db	2633	GCAGGAGTTGTTGAAATGAAATGTTGCTGGCAACAGCACTGAAAGATAAGAAAGAGCTA	2692
Qy	426	LysHisAlaGlyLysIleAlaThrGluAlaLeuGluAenIleArgThrIleValSerLeu	445
Db	2693	GAAGGAGCTGGGAAGATTGCTACAGAAGCCATCGAAAACTTCCGAACCTGTGTTCTTTG	2752
Qy	446	ThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsn	465
Db	2753	ACTCGGAGCAGAAAGTTTGAATACATGATGACAGAGTTTGCAGTACCACATACAGAAAC	2812
Qy	466	ThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyr	485
Db	2813	TCTTTGAGGAAGACACACATCTTCGGGGTCTCATTTTCTATCACCAGGCAATGATGAT	2872
Qy	486	PheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThr	505
Db	2873	TTTTCTCTATGCTGGGTGTTTCGGTTCGCTACTTGTGGTCAAAAGTGGTTCATGAAC	2932
Qy	506	ProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLys	525
Db	2933	TTTCAGGATGTTCTTTTGTGATTTCTCAGCTATGCTTTTGTGGTCCATGGCAGTGGGGCAG	2992
Qy	526	ThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAla	545
Db	2993	GTCAGTTTCATTTGCTCTGACTATGCTCAAGGCAAAAGTATACAGCAGCCCGCTCATCATG	3052
Qy	546	LeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThr	565


```
Db 3053 ATCATTGAAGAGCCCTCTGATTGACAGCTACAGCCCTCAGGCCCTCAGCCAAATACG 3112
QY 566 CyeGluGlyAenLeuGluPheAArgGluValSerPhePheTyrProCysArgProAspVal 585
Db 3113 TTGAAGGAAATGTGACATTTAATGAGTGTGTTCAACTATCCACTCCAGCAGACATC 3172
QY 586 PheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGly 605
Db 3173 CCCGTGCTCCAGGGCTGAGCCTCGAGTGAAGAGGCCAGAGCGCTGGCCCTCGTAGGT 3232
QY 606 SerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProVal 625
Db 3233 AGCAGTGGCTGTGGGAAGACACAGTTGTTGAGCTCTAGAGCGCTTCTATGACCCCTTG 3292
QY 626 GlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuValGlnTrpLeuArg 645
Db 3293 GCTGGTTTCAGTGCTAATTGATGCAAGAGATAAAGCACCTGAATGTCCAGTGGCTCCGA 3352
QY 646 SerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsn 665
Db 3353 GCACACCTGGGCATCGTGTCTCAGGAGCCCATCTGTTGACTGCAGCATTTGCCGAGAAC 3412
QY 666 IleAlaTyrGlyAspAenSerArgValValProLeuAspGluIleLysGluAlaAlaAen 685
Db 3413 ATTGCCTATGGAGACACACAGCCGGTCTATCATGTAAGAGATTGTGCAGGCGCAAG 3472
QY 686 AlaAlaAenIleHisSerPheIleGluGlyLeuProGluLysTyrAenThrGlnValGly 705
Db 3473 GAGGCCAACATACACACTTTCATCGACACTCCCTGAGAAATACAAACACAGAGTAGGA 3532
QY 706 LeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeu 725
Db 3533 GACAAAGGAACCCAGCTCTCTGTGGGCCAGAAACAGCGCATTCGCCATGTCCGCTCTT 3592
QY 726 LeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAenAspSer 745
Db 3593 GTTAGACAGCTCATATTTTGCTTTGGATGAGTACATCAGCTCTGGATCAGAAAGT 3652
QY 746 GluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValVal 765
Db 3653 GAAAGGTTGTCCAAGAGCCCTGGCAAGCCAGAGAGGCCGACCTGCTATTGTGATC 3712
QY 766 ThrHisArgLeuSerAlaIleGlnAenAlaAspIleValValLeuHisAenGlyLys 785
Db 3713 GCCCAGCGCTGTGCCACATCCAGATTCAGATTTAATAGTGTGTTTCAGAAATGCCAA 3772
QY 786 IleLysGlnGlnGlyThrHisGlnGluLeuArgAsnArgAspIleTyrPheLysLeu 805
Db 3773 GTCAGGAGCATGGCACATCAACAGCTGCTGGCCAGAGAGGCCGACCTATCTATTTTCCATG 3832
QY 806 ValAsnAlaGlnSer 810
Db 3833 GTCAGTGTCCAGGCT 3847
RESULT 11
AAZ52047
ID AAZ52047 standard; cDNA; 4369 BP.
XX
AC AAZ52047;
XX
DT 18-JUL-2000 (first entry)
XX
DE Rat multidrug resistance protein 1a cDNA.
XX
KW Multidrug resistance protein 1a; mdrla; multi-specific drug transporter;
KW drug formulation; formulation excipient; compound design; inflammation;
KW pharmacokinetic; oral absorption; cancer; cardiovascular disease;
KW central nervous system disorder; auto-immune disease; kidney disease; ss.
XX
OS Rattus rattus.
XX
FH Key Location/Qualifiers
FT CDS 352..4170
```

```
FT
FT
XX
PN WO200015650-A1.
XX
PD 23-MAR-2000.
XX
PF 10-SEP-1999; 99WO-US20770.
XX
PR 17-SEP-1998; 98US-0156800.
XX
PR 09-DEC-1998; 98US-0208809.
XX
PA (SMK ) SMITHKLINE BEECHAM CORP.
XX
PI Brun KA, Chenery RJ, Ellens H, Feild JA, Yue L;
XX
DR WPI: 2000-271372/23.
XX
DR P-PSDB; AAY70596.
XX
PT Isolated rat mdrla polynucleotides and polypeptides, useful in assays
PT to provide information on drug formulation, selection of formulation
PT excipients and compound design.
XX
PS Claim 2; Page 23-24; 33pp; English.
XX
CC The present cDNA sequence encodes rat multidrug resistance
CC protein 1a (mdrla). This protein is a member of the multi-specific drug
CC transporters family. Mdrla is used in assays to provide information on
CC drug formulation, selection of formulation excipients and compound
CC design. They are used in cell based, membrane based, binding or other
CC assays to provide information that may enhance drug formulation. This
CC invention further relates to the generation of in vivo and in vitro
CC comparison data to predict oral absorption and pharmacokinetics. This
CC enables the selection of drugs with optimal pharmacokinetics, i.e. good
CC oral bioavailability, brain penetration, plasma half life, and minimum
CC drug interaction. Transgenic and knock-out animals created using rat
CC mdrla provides an insight into treating and preventing human diseases
CC including cancer, inflammation, cardiovascular disease, central nervous
CC system disorders, auto-immune and kidney disease.
XX
SQ Sequence 4369 BP; 1241 A; 973 C; 1096 G; 1059 T; 0 other;
```

```
Alignment Scores:
Pred. No.: 2.18e-226 Length: 4369
Score: 2441.00 Matches: 470
Percent Similarity: 77.04% Conservative: 164
Best Local Similarity: 57.11% Mismatches: 175
Query Match: 59.84% Indels: 14
DB: 21 Gaps: 5
```

```
US-09-873-409-2 (1-812) x AAZ52047 (1-4369)
QY 2 ValAspGluAsnAspIleArgAlaLeuAenValArgHisTyrArgAspHisIleGlyVal 21
Db 1684 ATCGACGGACAGACATCAGGACCATCAATGTGAGTATCTCGGGAAATCATTTGGGGTG 1743
QY 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
Db 1744 GTGAGTCAGAAACCCGTGCTGTTGCCACCACCAATTCGCCGAAACATTCGCTATGSCCGA 1803
QY 42 AspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
Db 1804 GAAACGTCACCATGGATGAGTAAAGCTGTCAAGAAAGCCCAATGCCCTATGATTTTC 1863
QY 62 IleMetGluPheProAsnLysPheAenThrLeuValGlyLysGlyAlaGlnMetSer 81
Db 1864 ATCATGAAATCTCCCAAAATTTGACACCTTGTTGGTGGAGAGAGGGCGGACGCTGAGT 1923
QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
Db 1924 GGGGACAGAAACACAGAGGATGCCATTCGCCGGGCCCTGGTCCGCAACCCCAAGATCCTT 1983
QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 121
```


Db 1984 TTTGTTGGATGAGCCACGTCACCGCTTGGACACAGAAAGCCGCTGGTTCAGGCCGCT 2043
Qy 122 LeuGluLysAlaSerLysGlyArgThrThrIleValAlaHisArgLeuSerThrIle 141
Db 2044 CTGGATAGGCTAGAGAGCCGACCATTTGTATAGCTACCGCTGTGTACAGTT 2103
Qy 142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
Db 2104 CGCAATGCTGACGTCATTCGTTTGTATGCTGTGTCTCATTTGTGGAGCAAGAAATCAT 2163
Qy 162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln----- 178
Db 2164 GATGAGCTCATGAGAGAAAGAAATTTACTTCAAACTTGTCTGACTCAGACAGCAGGA 2223
Qy 179 ---AspIleLysAlaAspGluGlnMetGluSer-----MetThr 191
Db 2224 ATGAAATTTGAATTAGGAATGAGCTTGTGTAATCTAAGACGGAAATGTAATGTGGAC 2283
Qy 192 TyrSerThrGluArgLysThrAsnSerLeu-----ProLeuHisSerValLysSerIle 209
Db 2284 ATGTCTTCAAAAGATTCCGGATCCAGTCTTAATAAGAAAGAAATCAACTCGCAAAAGCATC 2343
Qy 210 LysSerAppPheIleAspLysAlaGluLysThr---GlnSerLysGluIleSerLeu 228
Db 2344 CGTGGGCCACATGATCAAGACCGGGAACCTTAGCACCAAGAGGCTCTGGATCAGCAGCTA 2403
Qy 229 ProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTyrProPheValVal 248
Db 2404 CTCACAGCTTCTTTTGGCGGATCCTGAAGTTGAATCAACTGAATGGCTTATTTTGTG 2463
Qy 249 LeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePhe 268
Db 2464 GTTGGTGTATTTTGTGCCATAATAATGGAGCTTGCAACACGACATCTCCATAATATT 2523
Qy 269 AlaLysIleIleThrMetPheGlyAsnAsnAspLysThrLeu---LysHisAspAla 287
Db 2524 TCAAAAGTTGTAGGGGTTTTTACAAAAATATGACACCCCTGAAATCCAGCGGCAAGACAGC 2583
Qy 288 GluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMet 307
Db 2584 AACTGTTTTCTTTATGTCTGTATCTTGGATCATCTTTTCAATACGTTTTTCTTT 2643
Qy 308 GlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla 327
Db 2644 CAGGCTTCACATTTGGCAAGCTGGAGAGATCCTCACCAAGGACTCCGATACATGTC 2703
Qy 328 PheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGluLysGluAsnSerThrGly 347
Db 2704 TTCAATCCATGCTGAGACAGACATAAGCTGTTTGTATGACCCCTAAAAACACCACAGGA 2763
Qy 348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyValThrGlySerArg 367
Db 2764 GCGCTGACCACCGGCTTGCCAAATGACGCTGCTCAAGTGAAGGGGCTACAGGGTCTAGG 2823
Qy 368 IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387
Db 2824 CTTCGCTGTATTTACCAGAACATAGCAATCTTTGGACAGGCAATCATATATCCCTGATC 2883
Qy 388 TyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGly 407
Db 2884 TACGGCTGGCAATTGACACTTTTACTCTTAGCAATTTGTTCCCATCTGCTATGATAGCAGGA 2943
Qy 408 MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnLysIleLysHis 427
Db 2944 GTGGTTGAAATGAAATGTTGTCTGGACACCGCTGAAAGATGAAGAAAGCAATAGAAAGT 3003
Qy 428 AlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArg 447
Db 3004 TCTGGAGAGATGCTACAGAGCAATTTGAATACTTTCGCACTGTCTCTTTTGACTCGG 3063
Qy 448 GlyLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSer 467

Db 3064 GAGCAGAAAGTTTGAACCTATGTATGCCAGAGCTTGCAGATACCATACAGAAATCCTTTG 3123
Qy 468 LysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487
Db 3124 AAGAAAGCGCAGCTTTTGGGATCACTTCTCTTCCACCGGCCATGATGATTTCTCC 3183
Qy 488 TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507
Db 3184 TATGCTGCTGTTTCCGTTTTCATGCTTACTTGTGTGGCAGCAAGAACTCATGACATTTGAA 3243
Qy 508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeu 527
Db 3244 AATGTTCTGTAGTATCTCAGCTATTTCTTTGTGTCATGGCAGTGGCGCGGTCACT 3303
Qy 528 ValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeu 547
Db 3304 TCATTCGCTCTGACTACGCGAAAGCAAGTCTCGGCATCCACATCATCAGGATCATTT 3363
Qy 548 GluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGlu 567
Db 3364 GAGAAATCCCTGAGATTACAGCTACAGCAGCGAGGGCTTGAAGCCTAATATGTTGAA 3423
Qy 568 GlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIle 587
Db 3424 GGAATGTGAATTTAATGGAGTCATGTTCAACTATCCACCCGACCAACATCCAGTG 3483
Qy 588 LeuArgGlyLeuSerLeuSerIleGluArgLysThrValAlaPheValGlySerSer 607
Db 3484 CTTCAGGGCTGAGCTAGAGGTGAAGAAAGGCGAGACGCTGCCCTCGTGGCGAGCAGT 3543
Qy 608 GlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGly 627
Db 3544 GGTCTCGGGAAGAGTACAGTGTCTGCTGAGTCTTATGATCCCATCGCCGGA 3603
Qy 628 GlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTyrLeuArgSerGln 647
Db 3604 ACAGTGTCTTAGTAGGCAAAAGAAATCAACTCAATGTCCAGTGGCTCCGCGCCAC 3663
Qy 648 IleAlaIleValProGlnProValLeuPheAsnCysSerIleAlaGluAsnIleAla 667
Db 3664 CTGGGCATTTGTCCAGAGGCCATCTCTTTGTACTGAGCATCGCCGAGAACATTGCC 3723
Qy 668 TyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAla 687
Db 3724 TACGAGACAAACAGCCGTGTCTGTCTATAAGAGAGATCGTGAAGCGACCCAGGAGGCC 3783
Qy 688 AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLys 707
Db 3784 AACATCCACCAAGTTTCATCGACTCACTGCTGAGAAATCAACACAGAGTGGGAGCAAAA 3843
Qy 708 GlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuGln 727
Db 3844 GGGACTCAGCTGTGCGGGCGGAGAGCGCATCGCCATCGCGCGGCCCTCTGTGAGA 3903
Qy 728 LysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAsnAspSerGluLys 747
Db 3904 CAGCTCATCTTACTTCTGGATGAAGCGACATCAGCTCTGATACGAGAGTGAAGAAG 3963
Qy 748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767
Db 3964 GTGCTCCAGGAAGCGCTGGCAAAAGCCAGGAGGCCGACCTGATTTGTGATCGCGCAC 4023
Qy 768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisGlnLysIleLys 787
Db 4024 GCGCTGTCCACCATCCAGACCGACATTTGATCGTGGTGAATTCAGAACCGCCGCGTCAAG 4083
Qy 788 GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807
Db 4084 GAGCAGCGCACCCACCACGAGCTGTGCGCCCAAGAAAGCATCTATTTCTCGATGTGCTCACT 4143
Qy 808 AlaGlnSer 810
Db 4144 GTGCAGGCT 4152

Db 2764 GCGCTGACACCGAGCTTGGCCAAATGACGCTGCTCAAGTCAAAAGGGGCTACAGGGTCTAGG 2823
Qy 368 IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387
Db 2824 CTTGCTGTATTACCCAGAACATGACAAATCTTGGGACAGGCATCATATCCCTGATC 2883
Qy 388 TyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGly 407
Db 2884 TAGGCTGGCAATTGACACTTTTACTCTAGCAATTGTTCCCATCATCTGCTATACAGGA 2943
Qy 408 MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHis 427
Db 2944 GTGGTGGAAATGAAATGTGTCTCGCAAGCGCTGAAAGATTAAGAAGAACTAGAAAGT 3003
Qy 428 AlaGlyLysIleAlaThrGluAlaLeuGluAsnIleAlaThrIleValSerLeuThrArg 447
Db 3004 TCTGGGAGATCGCTACAGAACATTTGAAAACCTTTCGCACCTGTCTCTTTGACTCGG 3063
Qy 448 GluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSer 467
Db 3064 GAGCAGAAAGTTGAAACTATGTATGCCAGAGCTTTCCTCTTCCAGAGCATCATGATGCTTTG 3123
Qy 468 LysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487
Db 3124 AAGAAAGCGCAGCTCTTTGGGATCACTTCTCTCTTCCAGGCGCATGATGATTTCTCC 3183
Qy 488 TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507
Db 3184 TATGCTGCTGTGTTCCGGTTTGATGCCCTACTTGTGGTGGCAGAGAACTATGATCATTTGAA 3243
Qy 508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeu 527
Db 3244 AATGTTCTGTAGTATTCTCAGCTATTGCTTTGTGGTGCATGGCAGTGGGCGAGTCACT 3303
Qy 528 ValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeu 547
Db 3304 TCATTGCTCTCTGACTAGCGGAAAGCCAAAGTCTCGGCATCCACATCATCAGGATCAT 3363
Qy 548 GluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGlu 567
Db 3364 GAGAAATCCCTTGAGATTGACAGCTACGACGAGGAGGCTTGAAGCCTTAATATGTTGAA 3423
Qy 568 GlyAsnLeuGluPheArgGluValSerPheThrProCysArgProAspValPheIle 587
Db 3424 GGAATGTGAAATTTAATGAGTCTATGTTCACTATCCACCCGACCAACATCCAGTG 3483
Qy 588 LeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSer 607
Db 3484 CTTCAAGGGCTGAGCTGAGGTGAAGAAAGGCGAGAGCTGGCCCTCTGTGGGCGAGCT 3543
Qy 608 GlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGly 627
Db 3544 GGCTCGGGAGAGTACAGTGTGTCAGCTGTGTGAGCGCTTCTATGACCCCATGGCCGGA 3603
Qy 628 GlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTyrLeuArgSerGln 647
Db 3604 ACAGTGTCTTAGTGGCAAGAAATTAAGCAACTCAATGTCCAGTGGTCCGCGCCAC 3663
Qy 648 IleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAla 667
Db 3664 CTGGGCATTGTGCTCCAGAGGCCATCTCTGTTGACTGCGAGCATCGCCGAGAACTTGC 3723
Qy 668 TyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAla 687
Db 3724 TACGAGACAAACAGCCGCTGCTGCTGCTTCAATAGAGAGATCGTAAGCGCAGCCAGGAG 3783
Qy 688 AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnGlnValGlyLeuLys 707
Db 3784 ACATCCACCACTTATCAGCTACCTGCTGAGAAATACACACAGAGTGGGAGACAAA 3843
Qy 708 GlyAlaGlnLeuSerGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGln 727

Db 3844 GGGACTCAGCTGTGCGGGCGGAGAGCAGCGCATCGCCATCGCGCGCCCTCTCTCAGA 3903
Qy 728 LysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 747
Db 3904 CAGCCTCACATCTTACTTCTGATGAACGACATCAGCTCTGGATACGAGAGTGAAGAAG 3963
Qy 748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767
Db 3964 GTCGTCCAGGAAGCGCTGGACAAAGCCAGGAAAGCCGACCTGCATTGTGATCGGCAC 4023
Qy 768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys 787
Db 4024 CCCTGTCCACATCCAGAACGACAGCTTGATGCTGGTGTATTCAGAACGCCAGGTCAAG 4083
Qy 788 GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807
Db 4084 GAGCAGCGCACCCACAGCAGCTGCTGCGCCAGAAAGGCATCTATTCTCGATGCTCAGT 4143
Qy 808 AlaGlnSer 810
Db 4144 GTGCAGGCT 4152
RESULT 13
RAD03489
ID AAD03489 standard; cDNA; 4317 BP.
XX
AC AAD03489;
XX
DT 13-JUN-2001 (first entry)
XX
DE Dog P-glycoprotein (PGP) cDNA #2.
XX
KW Dog; P-glycoprotein; PGP; multidrug transporter; MDRI;
KW drug bioavailability; transgenic animal; genetic model; ss.
XX
OS Canis familiaris.
XX
FH Key Location/Qualifiers
CDS 70..3912 /*tag= a /product= "Dog P-glycoprotein (PGP) #2"
FT
FT
XX WO200123540-A2.
XX
XX 05-APR-2001.
XX
XX 28-SEP-2000; 2000WO-US26767.
XX
XX 28-SEP-1999; 99US-0156510.
XX
XX (GENT-) GENTEST CORP.
XX
XX Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;
XX WPI; 2001-235373/24.
XX P-PSDB; AAE00304.
XX
XX New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
XX for determining the bioavailability of drugs and for screening for dog
XX PGP inhibitors -
XX
XX Claim 1; Page 66-72; 111pp; English.
XX
XX The invention relates to dog P-glycoprotein (PGP) also referred
XX as multidrug transporter (MDRI) and nucleic acids encoding them.
XX The invention also includes fragments and biologically functional
XX variants of dog P-glycoprotein. PGP and their nucleic acids are
XX useful for determining the bioavailability of drugs and for
XX screening PGP inhibitors. They are useful for the diagnosis and
XX treatment of conditions characterised by PGP activity, by
XX reducing or increasing PGP activity in a cell. PGP nucleic acids
XX are used as oligonucleotide probes. Complements of PGP nucleic
XX acids are useful as antisense oligonucleotides, to induce a PGP

'knockout' phenotype. They are used to prepare a non-human transgenic animal, which are valuable as genetic models for human diseases.
CC The present sequence is dog P-glycoprotein (pgp) cDNA. The
CC PGP enzyme functions as an efflux pump exporting small molecules
CC across the cell membrane. This enzyme is a member of the ABC
CC transporter family.
XX

SQ Sequence 4317 BP; 1293 A; 844 C; 1019 G; 1161 T; 0 other;

Alignment Scores:
Pred. No.: 4,18e-226 Length: 4317
Score: 2438.00 Matches: 467
Percent Similarity: 76.24% Conservative: 162
Best Local Similarity: 56.61% Mismatches: 178
Query Match: 59.77% Indels: 18
DB: 22 Gaps: 4

US-09-873-409-2 (1-812) x AAD03489 (1-4317)

Qy 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
Db 1429 ARTGATGGACAGACATTAGGACCAATAATGTAAAGCATCTTCGGGAATATTACTGGTGG 1488
Qy 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
Db 1489 GTGAGTCAGAGCGCTGTGTGTTGGCCACACGATAGCTGAAAAACATTCGCTATGGCCGC 1548
Qy 42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
Db 1549 GAAATGTCCACCATGATGAGATTGAGAAAGCTGTAAAGGAAGCCATGCTATGATTTT 1608
Qy 62 IleMetGlnPheProAsnLysPheAsnThrLeuValClyGluLysGlyAlaGlnMetSer 81
Db 1609 ATCATGAAACTACTATAATTTGACATCTGTTGGAGAGAGAGGGCCCGCTGAGT 1668
Qy 82 GlyGlyGlnLysGlnArgIleAlaAlaArgAlaLeuValArgAsnProLysIleLeu 101
Db 1669 GGTGGACAGAAACAGAGAAATCGCATTTGCTGGGCGCTGTTGCGCAACCCCAAGATCTT 1728
Qy 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValAlaAla 121
Db 1729 CTGCTGGATGAGCAACGTCAGCTCTGGACACTGAAAGTGAACGATGGTTCAGGTGGCC 1788
Qy 122 LeuGlyLeuAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
Db 1789 CTGATAAGCCAGAAAGCCGCGATACCATTTGTGATAGCTCATCTGTTGTCTACAGTT 1848
Qy 142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
Db 1849 CGTAATGCCGATGCTATGCTGTTTGTATGATGATGAGTCAATTTGGAGAAAGGAAATCAT 1908
Qy 162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIleLys 181
Db 1909 GATGAATCTCATGAAGAGAGGCGCATTTACTTCAAACTTGTCAATATGCAG---ACAAGA 1965
Qy 182 LysAlaAspGluGlnMetGluSerMetTyrSerThrGluArgLysThrAsnSerLeu 201
Db 1966 GGAATGAATGAGTAGTAAGAAATGCCATCTGGTGAATCCAAAGTGAAGTGAATGATGCTTG 2025
Qy 202 ProLeuHisSerValLysSerIleLysSerAspPheIle----- 214
Db 2026 GAAATG---TCTCCAAAAGATTCCAGGTCCAGTTTAAATAAAGAAAGATCAACTCGCAGG 2082
Qy 215 -----AspLysAlaGluSerThrGlnSerLysGluIle 226
Db 2083 AGTATACATGCCACACAGGCCAAGACAGAAAGCTTGGTACAAAGAGACTTGAATGAG 2142
Qy 227 SerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTyrProPhe 246
Db 2143 AATGTTCCCTCAGTTTCCTCTCGAGGATTTCTGAAGCTGAACCTCAACTGAATGCGCTTAT 2202
Qy 247 ValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIle 266

Db 2203 TTTGGTGGTGGATATATTTTGGCTATTATAAACGGAGGCTCCAACCCAGCATTTTCAATA 2262
Qy 267 IlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAspLysThrThrLeuLysHis 285
Db 2263 ATATTTTCAAGGATATATAGGATCTTTACCCGAGATGAGGATCTCTGAAACAAACACGACAG 2322
Qy 286 AspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyr 305
Db 2323 AATAGTAACATGTTTCTGTATTGTTCTAGTCTTGGAAATATTCTTTTATTATCATTT 2382
Qy 306 PheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHis 325
Db 2383 TTTCTTCCAGGGCTTCACATTTGGCAAGCTGGGAGATCTCTCACTAAGCGGCTTCGATAC 2442
Qy 326 LeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGluLysGluAsnSer 345
Db 2443 ATGGTTTTCAGATCCATGCTGAGACAGGATGTCAGCTGGTTTGATGACCCCTAAAAACACC 2502
Qy 346 ThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGly 365
Db 2503 ACTGAGCATTTGACAAACAGGCTTGCCAAATGATGGGCTCAAGTTTAAAGGGCTATAGT 2562
Qy 366 SerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSer 385
Db 2563 TCCAGGCTTGGCTGCTATTACCAGAAATATAGCAATCTTTGGGACAGGCATTATTATATCC 2622
Qy 386 PheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaVal 405
Db 2623 TTAATCTATGTTGGCAATTAACACTTTTACTTTAGCAATTTGACCATCATCTGCAATA 2682
Qy 406 ThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeu 425
Db 2683 GCAGGAGTGTGTTGAAATGAAATGTTGTCTGCAACAGCACTGAAAGATAGAAAGAGCTA 2742
Qy 426 LysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeu 445
Db 2743 GAAGAGCTGGGAAGATTTGCTACAGAGCCATCGAAATCTTCGAACTGTTGTTCTTTG 2802
Qy 446 ThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsn 465
Db 2803 ACTCGGACAGCAAGTTTGAATACATGATGACAGAGTTTGCNAAGTACCATACAGAAAC 2862
Qy 466 ThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyr 485
Db 2863 TCTTTGAGGAAAGACACATCTTCGGGGTCTCATTTTCTATCACCAGGCAATGATGAT 2922
Qy 486 PheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThr 505
Db 2923 TTTTCTATGCTGGCTGTTTTCGGTTCGCTACTTGTGGCAAAATGAGTTCATGAAC 2982
Qy 506 ProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLys 525
Db 2983 TTTGAGGATTTCTTTTGGTATTTCTCAGCTATTTGTTGGTGGCCATGCGATGGGGCAG 3042
Qy 526 ThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAla 545
Db 3043 GTCAGTTTCATTTGCTCTGACTATGCCAAAGCAAAAGTATCAGCAGCCCAACGTCATCATG 3102
Qy 546 LeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThr 565
Db 3103 ATCATTTGAAAGAAAGCCCTCTGATTTGACAGTACAGCCCTCAGCGGCTCAAGCCAAATACG 3162
Qy 566 CysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspVal 585
Db 3163 TTTGAGGAAATGTGACATTTAATGAGGTCGTCTTCAACTATCCCATCGACAGACATC 3222
Qy 586 PheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGly 605
Db 3223 CCGTCTCTCCAGGGGCTGAGCTCGAGGCTGAAAGAGGGCCAGACGCTGGCCCTCGTAGGT 3282
Qy 606 SerSerGlyCysGlyLysSerThrSerValGlnLeuGluArgLeuTyrAspProVal 625


```
Db 3283 AGCAGTGGCTGTGGGAAGACACAGTGTTCAGCTCCTTAGACGGCTTCTATGACCCCTTG 3342
Qy 626 GlnGlyGlnValLeuPheAspGlyValAlaLysGluLeuAsnValGlnTrpLeuArg 645
Db 3343 GCTGGTTCAGTCTAATTGATGGCAAGAGATAAGACCACTGAATGTCAGTGGCTCGA 3402
Qy 646 SerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsn 665
Db 3403 GCACACCTGGGCATCGTGTCTCAGGAGCCCATCTGTGTGACTCGAGCATTCGCCGAGAAC 3462
Qy 666 IleAlaTyrglyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsn 685
Db 3463 AFTGCTTATGGAGACAACAGCCGGTCTGATCACATGAAGAGATTATGAGGAGCCCAAG 3522
Qy 686 AlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyraAsnThrGlnValGly 705
Db 3523 GAGGCCACATACACCACTTCATCGAGACACTCCCTGAGAAATACACACAGATGAGA 3582
Qy 706 LeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeu 725
Db 3583 GACAAAGGAACCCAGCTCTCTGTGGCCAGAAACAGCGCATTGCCATAGCTCGCGCTCT 3642
Qy 726 LeuGlnLysProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspAsnSer 745
Db 3643 GTTAGACAGCCCTCATATTTGCTTTTGGATGAAGCTACATCAGCTCTGGATACAGAAAT 3702
Qy 746 GluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValVal 765
Db 3703 GAAAGTTGTCCAGACCCCTGNCNAAGCCAGAGAGGCGCACCTGCGATTGTGATC 3762
Qy 766 ThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLys 785
Db 3763 GCCCACCCTGTGCCACCATCCAGAATGCAGATTTAATAGTGTGTTTCAGAAATGGCAAA 3822
Qy 786 IleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrrPheLysLeu 805
Db 3823 GTCAAGGAGCATGGCACATCAACAGCTGCTGGCTCAGAAAGGCGCATCTATTTTCCATG 3882
Qy 806 ValAsnAlaGlnSer 810
Db 3883 ATCAGTGTCCAGCT 3897
RESULT 14
ABL91687
ID ABL91687 standard; DNA; 3840 BP.
XX ABL91687;
AC ABL91687;
XX
XX 28-MAY-2002 (first entry)
XX Human polynucleotide SEQ ID NO 30.
XX
XX Human; HIV; HCV; gene expression; oligoribonucleotide; tumour; pathogen;
KW Plasmidium; virus; viroid; cytokine; prion; antisense oligonucleotide;
KW cytostatic; virucide; protozoacide; antibacterial; ds.
XX
XX Homo sapiens.
XX
XX DE10100586-C1.
XX
XX 11-APR-2002.
XX
XX 09-JAN-2001; 2001DE-1000586.
XX
XX 09-JAN-2001; 2001DE-1000586.
XX (RIBO-) RIBOPHARMA AG.
XX
XX Kreutzer R, Limmer S, Rost S, Hadwiger P;
XX
XX WPI; 2002-270454/32.
XX
XX Inhibiting gene expression in cells, useful for e.g. treating tumors,
PT
```

by introducing double-stranded complementary oligoRNA having unpaired terminal bases -

Claim 13; Page 28-30; 104pp; German.

The invention relates to a method for inhibiting expression of a target gene (ABL91687-ABL91797) in a cell by introducing at least one oligoribonucleotide that has a double-stranded structure consisting of at most 49 sequential nucleotide pairs, with at least part of one strand complementary with the target gene and has at least one end a single-stranded segment of 1-4 nt. The method provides oligoribonucleotides for antisense inhibition of gene expression useful e.g. for treating tumours but the oligoribonucleotides may also be directed against genes present in pathogens (e.g. Plasmidium or cyruses/viroids, pathogenic on humans, animals or plants) or against cytokine, id, developmental or prion genes. The method provides more effective inhibition of gene expression than use of known oligonucleotides, probably because the unpaired overhang increases stability and thus intracellular concentration.

Sequence 3840 BP; 1130 A; 738 C; 957 G; 1015 T; 0 other;

Alignment Scores:

Pred. No.:	5.51e-226	Length:	3840
Score:	2436.00	Matches:	466
Percent Similarity:	76.79%	Conservative:	166
Best Local Similarity:	56.62%	Mismatches:	177
Query Match:	59.72%	Indels:	14
DB:	24	Gaps:	4

US-09-873-409-2 (1-812) x ABL91687 (1-3840)

Qy 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyraAspHisIleGlyVal 21

Db 1354 GTTGATGGACAGGATATTAGGACCATTAATGTAGGTTTCTACGGGAATCATTTGGTGG 1413

Qy 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrglyArg 41

Db 1414 GTGAGTCAGGAACCTGTATTGTTGCCACCATAGCTGAAAACATTCGCTATGCCCT 1473

Qy 42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyraPhe 61

Db 1474 GAAATGTGCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCCAATGCCATGACTTT 1533

Qy 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81

Db 1534 ATCATGAATACTGCCTCATTAATTTGACACCTGTTGGAGAGAGAGGGCCCGGTTGAGT 1593

Qy 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101

Db 1594 GGTGGGCAAGACAGAGGATCGCCATTGACCGCTGGTTCGCAACCCCAAGATCCTC 1653

Qy 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121

Db 1654 CTGCTGGATGAGGCCACGTCAGCTTGGACACAGAAAGCAAGCAGTGGTTTCAGTGGCT 1713

Qy 122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141

Db 1714 CTGGATAGCCCAAGAAAGTTCGGACCATTTGTGATAGCTCATCGTTGTCTACAGTT 1773

Qy 142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161

Db 1774 CGTAATGCTGACGTCATCGCTGTTTCGATGATGAGTCAATTCGAGAAAGGAAATCAT 1833

Qy 162 AlaGluLeuMetAlaLysArgGlyLeuTyrrTyrrSerLeuValMetSerGln----- 178

Db 1834 GATGAACATCAAGAGAAAGGCACTTACTTCAAACTTGTCAATATGCAGACAGCAGGA 1893

Qy 179 -----AspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrrSerThrGlu 195

Db 1894 ATGAAGTTGAATTGAATAATGACGCTGATGAATCCAAAGTGAAATTCATGCTTGGAA 1953

Qy 196 ArgLysThrAsnSerLeuProLeuHisSerVal-----LysSerIle 209


```
Db 1954 ATGCTCTCAAAATGATTCAAGATCCAGTCTAAATAAGAAAAAGATCAACTCGTAGGAGTGC 2013
Qy 210 Lys--SerAspPheIleApsLysAlaGluSerThrClnSerLysGluIleSerLeu 228
Db 2014 CDTGGATCAAGCCCAAGACAGAAAGCTAGTACCAAGAGGCTGTGAGAAATATA 2073
Qy 229 ProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValVal 248
Db 2074 CTTCCAGTTTCTTTTGGAGGATATGAAGCTAAATTTAACTGAATGGCTTATTTTGT 2133
Qy 249 LeuGlyThrLeuAlaSerValLeuAenGlyThrValHisProValPheSerIleIlePhe 268
Db 2134 GTTGGTGTATTTGTGCCATTATAAATGGAGGCTGCAACACGACCAITTTGCAATAATATT 2193
Qy 269 AlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeuLysHisAspAla 287
Db 2194 TCAAGATTATAGGGGTTTTTACAAGAATTTGATGATCCTGAAACAAACGACAGAAATAGT 2253
Qy 288 GluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMet 307
Db 2254 AACTGTTTTCATATGTTTCTAGCCCTTGGAAATATTTCTTTTATACATTTTCCCT 2313
Qy 308 GlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla 327
Db 2314 CAGGGTTTCACATTTGGCAAGCTGGAGAGATCCTCACCAGCGGCTCCGATACATGTTT 2373
Qy 328 PheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspLysGluAsnSerThrGly 347
Db 2374 TTTCCGATCCATGCTACAGACAGGATGTAGTGTGTTGATGACCTTAAACACACCACTGGA 2433
Qy 348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg 367
Db 2434 GCATTGACTACAGGCTCGCAATGATGCTGCTCAAGTTAAAGGGCTATAGTTCCAGS 2493
Qy 368 IleGlyValLeuThrGlnAenAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387
Db 2494 CTTGCTGTAATTACCCAGATATAGCAAACTTTGGGACAGGAATAATATATCTCTCATC 2553
Qy 388 TyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGly 407
Db 2554 TATGGTTGGCAACTAACTACTTACTCTTAGCAATTTGTACCCATCATTTGCAATAGCAGGA 2613
Qy 408 MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHis 427
Db 2614 GTTGTGGAATGAAATGTTGCTGGACAGCACTGAAAGATAGAAAGACTAGAGGT 2673
Qy 428 AlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArg 447
Db 2674 GCTGGGAAGATCGCTACTGAAGCAATAGAAAACCTCCGAACCGTTGTTCTTTGACTCAG 2733
Qy 448 GluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAenThrSer 467
Db 2734 GAGCAGAAGTTTGAACATATGTATGTCTCAGAGTTTGCAGGTACCATACAGAAACTCTTTG 2793
Qy 468 LysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487
Db 2794 AGGAAGCACACATCTTTGGAATTTACATTTTCTTACCAGCAATGATGATTTTTC 2853
Qy 488 TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507
Db 2854 TATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGCCACATAAATCATGAGCTTTGAG 2913
Qy 508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeu 527
Db 2914 GATGTTCTGTAGTATTTTCAGCTGTTGCTTTGGTGCCATGGCCGTGGGCGCAAGTCAGT 2973
Qy 528 ValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeu 547
Db 2974 TCATTTGCTCTGACTATGCAAGCAAGCAAAATATATCAGCAGCCCCACATCATCATCAT 3033
Qy 548 GluLysLysProAenIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGlu 567
```

```
Db 3034 GAAAAAACCCTTTTGATTGACAGCTACAGCAGGAGGCTTAATGCCGAAACATTGGAA 3093
Qy 568 GlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIle 587
Db 3094 GGAATGTACATTTTGGTGAAGTTGTATTTCACTATCCACCCGACCGACATCCAGTG 3153
Qy 588 LeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSer 607
Db 3154 CTTAGGAGACTCAGCTGGAGGTGAAGAGGCCAGACGCTGGCTCTGTGTGGCGACAGT 3213
Qy 608 GlyCysGlyLysSerThrSerValGlnLeuGlnArgLeuTyrAspProValGlnGly 627
Db 3214 GCGTGTGGGAAGAGACACAGTGTCTCCAGCTCTCCGGAGCGGTTCTACGACCCCTTGGCAGG 3273
Qy 628 GlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGln 647
Db 3274 AAAGTGCTCTGTATGGCAAGAAATAAAGCAGCTGAATGTTCAGTGGCTCCGAGCACAC 3333
Qy 648 IleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAla 667
Db 3334 CTGGGATCGTCTCCAGGAGCCCATCTGTTTGACTGCAGCATTTGCTGAGAACAATTGCC 3393
Qy 668 TyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAla 687
Db 3394 TATGAGACAAACAGCGGGTGTGTCTACAGGAAGAGATTGTGAGGCGCAGCAAGGAGGCC 3453
Qy 688 AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLys 707
Db 3454 AACATACATCGCTTATCGAGTCACTGCTCTAATAATATAGCATAAAGTAGGAGACAA 3513
Qy 708 GlyAlaGlnLeuSerGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGln 727
Db 3514 CGAACTCAGCTCTGTGTGGCCAGAAAACAAACGCAATTCCTAGCTCGTCCCTTGTAGA 3573
Qy 728 LysProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 747
Db 3574 CAGCCTCATATTTTGTGTTGATGAAGCCACGTCAGCTCTGGATACAGAAAGTAAAG 3633
Qy 748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767
Db 3634 GTTGTCCAAGAGCCCTGGCAAGCCAGAGAGCCGACCTGATTTGATTGTTGCTCAC 3693
Qy 768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys 787
Db 3694 CGCTGTCCACCATCCAGAAATGCAGACTTAATAGTGTGTTTTCAGAATGCGCAGTCAAG 3753
Qy 788 GluGlnGlyThrHisGlnLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807
Db 3754 GAGCATGGCAGCATCAGCAGTGTGTCGACAGAAAGGCATCTATTTTCAATGGTCACT 3813
Qy 808 AlaGlnSer 810
Db 3814 GTCCAGGCT 3822
RESULT 15
AAZ49332
ID AAZ49332 standard; cdna; 3860 BP.
XX AAZ49332;
AC AAZ49332;
XX
DT 14-MAR-2000 (first entry)
XX
DE Human wild-type multidrug resistance-1 (MDR-1) cDNA.
XX
KW Multidrug resistance; MDR-1; P-glycoprotein;
KW transmembrane efflux pump; haematopoietic stem cell; transduction;
KW bone marrow transplantation; chemotherapy; radiation therapy; cancer;
KW gene therapy; gene replacement; genetic defect; thalassaemia;
KW Gaucher's disease; sickle cell anaemia; leukaemia; ex vivo expansion;
XX cytokine; wild-type; ds.
OS Homo sapiens.
XX
```


FH Key Location/Qualifiers
 FT CDS 1..3843
 FT /*tag= a
 FT /product= "Human wild-type MDR-1 protein"
 FT mutation
 FT /tag= b
 FT /note= "cDNA sequence of G185V human mutant MDR-1 given
 in AA249333"
 XX
 PN W09961589-A2.
 XX
 XX 02-DEC-1999.
 XX
 PP 27-MAY-1999; 99WO-US11825.
 XX
 XX 28-MAY-1998; 98US-0086988.
 XX
 PA (SUJD-) ST JUDE CHILDREN'S RES HOSPITAL.
 XX
 XX Sorrentino B, Bunting K;
 PI
 DR WPI; 2000-072615/06.
 DR P-PSDB; AAY58186.
 XX
 XX Ex vivo expansion of hematopoietic stem cells transduced with a
 PT sequence encoding human multidrug resistance-1, used for bone marrow
 PT transplantation -
 PS
 PS Claim 10; Page 68-70; 113pp; English.
 XX
 CC This sequence represents cDNA encoding human wild-type
 CC multidrug resistance protein MDR-1. MDR-1 is a transmembrane
 CC efflux pump, responsible for the export of drugs from cells,
 CC particularly cancer cells. Wild-type MDR-1 shows increased
 CC resistance to etoposide and decreased resistance to vinca
 CC alkaloids compared with a mutant form (AAY58187) where the Gly at
 CC position 185 is replaced by Val. The invention relates to transducing
 CC hematopoietic stem cells with nucleic acid encoding an MDR protein
 CC and culturing the modified cells. The modified hematopoietic stem
 CC cells are useful in bone marrow transplantation (to reconstitute
 CC hematopoietic systems in patients who have undergone chemotherapy or
 CC radiation therapy) and in ex vivo gene therapy of genetic defects in
 CC cells derived from hematopoietic stem cells, e.g., thalassemia,
 CC Gaucher's disease, sickle cell anaemia or leukaemia. The modified
 CC cells can also be used to identify factors involved in regulating
 CC proliferation and differentiation in hematopoietic stem cells.
 CC Hematopoietic stem cells that express MDR-1 will be protected against
 CC chemotherapeutic agents, so can be engrafted while the patient is
 CC undergoing chemotherapy. Expansion of (rare) hematopoietic stem cells
 CC provides sufficient cells to permit standard biochemical analysis.
 CC Overexpression of MDR-1 allows cytokine-driven expansion of
 CC hematopoietic stem cells by at least 10-fold compared with a maximum
 CC of 4-fold in known procedures.
 XX
 SQ Sequence 3860 BP; 1135 A; 746 C; 957 G; 1022 T; 0 other;
 Alignment Scores:
 Pred. No.: 5,558-226 Length: 3860
 Score: 2436.00 Matches: 466
 Percent Similarity: 76.79% Conservative: 166
 Best Local Similarity: 56.62% Mismatches: 177
 Query Match: 59.72% Indels: 14
 DB: 21 Gaps: 4
 US-09-873-409-2 (1-812) x AA249332 (1-3860)
 Qy 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
 Db 1357 GTTGATGGACAGATATTAGGACCATTAATGTAAGGTTCTACGGGAATCATTTGGTGTG 1416
 Qy 22 ValSerGlnGluProValLeuPheGlyThrIleSerAsnAsnIleLysTyrGlyArg 41
 Db 1417 GTGAGTCAGGAACCTGTATTGTTGGCCACCAGTAGCTGAAACACATTCGCTATGCCCGT 1476

QY 42 AspAspValThrAspGluMetGluArgAlaAlaAa-gGluAaAsnAlaTyrAspPhe 61
 Db 1477 GAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAATGCCTATGACTTT 1536
 QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyGlyValAlaGlnMetSer 81
 Db 1537 ATCATGAACATGCCTCTATAAATTTGACACCCCTGGTGGAGAGAGAGGGGCCAGTTGAGT 1596
 QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
 Db 1597 GGTGGGAGAGAGCAGAGGATCGCCATTGCACTGCGCTGGTTCGCAACCCCAAGATCCTC 1656
 QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
 Db 1657 CTGCTGGATAGGCCACGTCAGCTTGACACAGAAAGCGAAGCAGTGGTTCCAGGTGGCT 1716
 QY 122 LeuGlyLysAlaSerLysGlyArgThrIleValValAlaAlaHisArgLeuSerThrIle 141
 Db 1717 CTGGATAAGGCCAGAAAGGTGGACCACTATTGTGATGCTCATCGTTGTCTACAGTT 1776
 QY 142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlyLysGlyAlaHis 161
 Db 1777 CGTAATGCTGACGTCATCGCTGGTTTCGATGATGAGTCATTGTGAGAAAGAAATCAT 1836
 QY 162 AlaGluMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln----- 178
 Db 1837 GATGAACCTCATGAAGAGAAAGGCATTTACTTCAAACTGTTCACAAATGCAGACAGCAGGA 1896
 QY 179 -----AspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu 195
 Db 1897 AATGAAGTTGAATTAGAAATGCAGCTGATGAATCAAAAGTGAAATGATCCCTGGAA 1956
 QY 196 ArgLysThrAsnSerLeuProLeuHisSerVal-----LysSerIle 209
 Db 1957 ATGCTTCAATGATTCAAGATCCAGTCTAATPAGAAAGATCAACTCGTAGGAGTGC 2016
 QY 210 Lys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLysGluSerLeu 228
 Db 2017 CGTGATCACAAGCCCAAGACAGCAAGAGCTTAGTACAAAGAGGCTCTGGATGAAAGTATA 2076
 QY 229 ProGluValSerLeuLysIleLeuLysLeuAsnLysProGluTyrProPheValVal 248
 Db 2077 CTTCCAGTTTCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGGCTTTATTTGTT 2136
 QY 249 LeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePhe 268
 Db 2137 GTTGGTGTATTTTGGCCATTATAATGGAGCCTGCAACCAAGCATTTGCAATATATTT 2196
 QY 269 AlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeuLysHisAspAla 287
 Db 2197 TCAAGATTATAGGGGTTTTCACAGAATTGATGATCCTGAAACAAACAGCAGATACT 2256
 QY 288 GluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMet 307
 Db 2257 AACTTGTGTTTCACTATTGTTCTAGCCCTTGGAAATATTCTTTTATTACATTTTCTCT 2316
 QY 308 GlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla 327
 Db 2317 CAAGGTTTCACATTTGGCAAGCTGGAGATCCTCACCAGCGGCTCGGATACATGTT 2376
 QY 328 PheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGlyLysGluAsnSerThrGly 347
 Db 2377 TTCCGATCCATGCTCAGACAGGATGTGAGTTGTTGATGACCTTAAACACCACTGGA 2436
 QY 348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg 367
 Db 2437 GCATTGACTACCAAGCTGCCCAATGCTGCTCAAGTTAAAGGGGCTATAGGTTCCAG 2496
 QY 368 IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387
 Db 2497 CTTGCTGTAATTAACCAAGATATACCAAACTCTGGGACAGGAATAATATATATCTTCATC 2556


```
QY 388 TyrGlyTrpGluMetThrPheLeuLeuLeuSerIleAlaProValLeuAlaValThrGly 407
DB 2557 TATGGTTGGCAACTAACACTGTTACTCTTAGCAATTTGTACCCATCATTTGCAATAGCAGCA 2616
QY 408 MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHis 427
DB 2617 GTTGTTCGAATGAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAAGAACTAGAAAGT 2676
QY 428 AlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArg 447
DB 2677 GCTGGGAAGATCCCTACTGAAGCAATAGAAAACCTCCGAACCGTTGTTCTTTGACTCAG 2736
QY 448 GluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSer 467
DB 2737 GACGAGAAGTTTGAACATATGATGTCAGAGTTTGCAGGTACCATACAGAAACTCTTTG 2796
QY 468 LysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487
DB 2797 AGGAAAGCACACATCTTTGGAAATTTACATTTCTCCAGGCAATGATGATATTTTCC 2856
QY 488 TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507
DB 2857 TATGCTGGATGTTCCGGTTTGGAGCTACTTGTGGCACATAAACTCATGAGCTTTGAG 2916
QY 508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeu 527
DB 2917 GATGTTCTGTAGTATTTTTCAGCTGTTGCTTTGGTGCCATGCGCCGTGGGCAAGTCAGT 2976
QY 528 ValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeu 547
DB 2977 TCATTTTGCTCCTGACTATGCAAGCAAGCAAAATATACAGCAGCCACATCATCATGATCAT 3036
QY 548 GluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGlu 567
DB 3037 GAAAAAACCCCTTTGATTTGACACTACAGCAGCAAGGCGCTAATGCCGAAACATTTGGAA 3096
QY 568 GlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIle 587
DB 3097 GGAATGTCACATTTGGTGAAGTTGTATTCAACTATCCACCCGACCGGACATCCCGAGTG 3156
QY 588 LeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSer 607
DB 3157 CTTTCAGGGAAGTGGAGCTGGAGGTGAAGAGGCGGCACAGCTGGCTCTGTGGGGCAGCAGT 3216
QY 608 GlyCysGlyLysSerThrSerValGlnLeuGlnArgLeuTyrAspProValGlnGly 627
DB 3217 GGCCTGGGAAGAGCACAGTGGTCCAGCTCTGGAGCGGTTCTACGACCCCTTTGGCAGGG 3276
QY 628 GlnValLeuPheAspGlyValAlaAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGln 647
DB 3277 AAAGTGTGCTTGTATGGCAAGAAATAAAGCGACTGAATGTTTCAGTGGCTCCGAGCACAC 3336
QY 648 IleAlaIleValProGlnGluProValPheAsnCysSerIleAlaGluAsnIleAla 667
DB 3337 CTGGCATCGTGTCCAGGAGCCCATCTCTTTGACTGCACGATTTGCTGAGAACATTTGCC 3396
QY 668 TyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAla 687
DB 3397 TATGGAGACAAACAGCCGGGGTGTCTACAGGAAGAGATCTGTAGGGGCGGCAAGAGGGCC 3456
QY 688 AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLys 707
DB 3457 AACATACATGCCCTTCATCGAGTCACTGCCTTAATAATATAGCACTAAAGTAGGAGACAA 3516
QY 708 GlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGln 727
DB 3517 GGAACCTCAGCTCTCTGGTGGCCAGAAACACGCAATGCCATAGCTCGTCCCTTGTAGA 3576
QY 728 LysProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 747
DB 3577 CAGCCCTCATATTTTGGATTTGGATGAAGCCAGCGTCAGCTCTGGATACAGAAAGTGAAG 3636
QY 748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767
```

```
DB 3637 GTTGTCCAAAGAGCCCTGGACAAAGCCGAGAGAGCCGACCTGCTATTGTTGCTCAC 3696
QY 768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValLeuHisAsnGlyLysIleLys 787
DB 3697 CCGCTGTCCACCATCCAGATGCAGACTTAATAGTGGTGTTCAGATGGCAGAGTCAAG 3756
QY 788 GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807
DB 3757 GAGCATGGCAGCCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTTCAATGGTCACT 3816
QY 808 AlaGlnSer 810
DB 3817 GTCCAGGCT 3825
```

Search completed: March 30, 2003, 03:08:04
Job time : 937.594 secs

GenCore version 5.1.4 p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 30, 2003, 03:22:44 ; Search time 5585.08 Seconds
(without alignments)
2354.618 Million cell updates/sec

Title: US-09-873-409-2
Perfect score: 4079
Sequence: 1 MVDENDIRALNVRHYRDHIG.....QELLNRDIYFKLVNAQSVQ 812

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n model -DRV=xlh
-Q/cgn2_1/USPTO_spool/US09873409/runat_27032003_115420_19240/app_query.fasta_1.7544
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09873409 @CCN 1 1 30544 @runat 27032003_115420_19240 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hcti:*
9: gb_est1:*
10: gb_est2:*
11: gb_hcti:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_prt:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1008	24.7	943	9	AL520322	AL520322 AL520322
2	834	20.4	998	14	BM904842	BM904842 AGENCOURT
3	780	19.1	2676	11	AK014319	AK014319 Mus muscu
4	770.5	18.9	1019	12	BG248052	BG248052 602359987
5	753	18.5	726	12	BG293345	BG293345 602390738
6	751	18.4	944	12	BF796582	BF796582 602358463
7	743.5	18.2	886	17	BH139685	BH139685 ENTNA47TR
8	740.5	18.2	871	17	AZ682350	AZ682350 ENTXB16TF
9	734.5	18.0	932	17	AZ670821	AZ670821 ENTJN69TF
10	728.5	17.9	947	17	AZ683753	AZ683753 ENTIL96TF
11	726.5	17.8	886	17	AZ540627	AZ540627 ENTQ18TF
12	723.5	17.7	880	17	AZ687805	AZ687805 ENTJU52TF
13	721.5	17.7	939	14	BQ720763	BQ720763 AGENCOURT
14	720.5	17.7	897	17	AZ541090	AZ541090 ENTDS67TR
15	711.5	17.4	913	17	BH155700	BH155700 ENTRO54TR
16	709.5	17.4	834	17	AZ548312	AZ548312 ENTG07TR
17	708.5	17.4	853	17	AZ679807	AZ679807 ENTH167TR
18	705.5	17.3	823	17	AZ532602	AZ532602 ENTCSR50TR
19	697.5	17.1	861	14	BQ717101	BQ717101 AGENCOURT
20	691	16.9	780	12	BG587938	BG587938 EST489713
21	690.5	16.9	899	17	BH154857	BH154857 ENTRO47TF
22	689.5	16.9	891	17	AZ682250	AZ682250 ENTK168TR
23	688	16.9	1123	11	AY108485	AY108485 Zea mays
24	683.5	16.8	1341	11	AY108285	AY108285 Zea mays
25	683	16.7	600	13	BM486593	BM486593 pgm2n.pk0
26	682.5	16.7	919	17	AZ690701	AZ690701 ENTIV68TR
27	678	16.6	926	14	BQ123477	BQ123477 EST609053
28	677.5	16.6	795	12	BF313560	BF313560 601900192
29	669	16.4	547	13	BQ307071	BQ307071 BJ030711
30	667	16.4	937	17	BH150760	BH150760 ENTQ042TR
31	663	16.3	835	13	BQ349604	BQ349604 BJ349604
32	662	16.2	636	10	AV962688	AV962688 AV962688
33	661	16.2	747	13	BQ445774	BQ445774 BJ445774
34	659.5	16.2	628	12	BQ080311	BQ080311 H3052B06-
35	659	16.2	852	17	BH720383	BH720383 BOH2294TF
36	657	16.1	715	14	BQ689512	BQ689512 QGD6G09.Y
37	657	16.1	785	10	AV709991	AV709991 AV709991
38	655.5	16.1	787	12	BG584063	BG584063 EST485823
39	655	16.1	609	10	BQ724901	BQ724901 AGENCOURT
40	653	16.0	609	10	AV986144	AV986144 AV986144
41	651.5	16.0	921	17	AZ687628	AZ687628 ENTIR96TR
42	651	16.0	746	13	BQ380617	BQ380617 BJ380617
43	650	15.9	729	13	BQ353293	BQ353293 BJ353293
44	647	15.9	1813	11	U66688	U66688 Homo sapien
45	641	15.7	820	13	BJ356540	BJ356540 BJ356540

ALIGNMENTS

RESULT 1
AL520322
LOCUS
DEFINITION AL520322 LTI_NFL004_NBC2 Homo sapiens cDNA clone CS0DB006YC15 5
prime, mRNA sequence.
ACCESSION AL520322
VERSION AL520322.1 GI:12783815
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 943)
AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization

JOURNAL COMMENT

Unpublished (2001)
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
 Location/Qualifiers
 1. .943

FEATURES

source
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CS08006YC15"
 /clone_lib="LTI_NFL004_NBC2"
 /sex="male"
 /tissue_type="neuroblastoma cells"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 253 a 225 c 253 g 211 t 1 others
 ORIGIN

Alignment Scores:

Pred. No.: 8.4e-107 Length: 943
 Score: 1008.00 Matches: 194
 Percent Similarity: 80.40% Conservative: 48
 Best Local Similarity: 64.45% Mismatches: 58
 Query Match: 24.71% Indels: 1
 DB: 9 Gaps: 0

US-09-873-409-2 (1-812) x AL520322 (1-943)

Qy 511 ileValPheThrAlaIleAlaTyGlyValAlaMetAlaIleGlyLysThrLeuValLeuAla 530
 Db 5 TTATATTTCAGCTGTGTCTTTGGTGCCATGCGCGGGGCAAGTCAGTTGCT 64
 Qy 531 ProGluTySerLysAlaLysSerGlyValaAlaHisLeuPheAlaLeuLeuGluLysLys 550
 Db 65 CTGACTATGCGCAAGCAAAATATACAGCAGCCACATCATCATGATGATGAAAAAAC 124
 Qy 551 ProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeu 570
 Db 125 CCTTTGATTGACAGCTACAGCAGCGAAGGCGCTAATGCCGACACATTGGAGGAAATGTC 184
 Qy 571 GluPheArgGluValSerPhePheTyProCysArgProAspValPheIleLeuArgGly 590
 Db 185 ACATTTGGTGAAGTTGTATTCACTATCCACCCGACCGACATCCAGTGCTTCAGGGA 244
 Qy 591 LeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGly 610
 Db 245 CTGAGCTGGAGGTGAGAGAGGGCCAGACGCTGCTGTGGTGGCAGCAGTGCTGTGGG 304
 Qy 611 LysSerThrSerValGlnLeuLeuGlnArgLeuTyTrpAspProValGlnGlnValLeu 630
 Db 305 AAGAGCAGTGGTCCAGCTCTGGAGCGGTCTACGACCCCTTTGGCAGGAGAAATGCTG 364
 Qy 631 PheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIle 650
 Db 365 CTTGATGCGCAAGAAATAAAGCGATGTAATGTTGAGTGCTCCGAGCAGCCTGGGCATC 424
 Qy 651 ValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyGlyAsp 670
 Db 425 GTGTCCAGGAGCCCATCTCTTTGAGCTGACAGCTGCTGAGAACATTGCTATGGAGAC 484
 Qy 671 AsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHis 690
 Db 485 AACAGCCGGGTGGTGTCTACAGGAAGAGATCGTGGGGCAGCAAAAGAGGCCCAACATCAT 544

Qy 691 SerPheIleGluGlyLeuProGluLysTyAsnThrGlnValGlyLeuLysGlyAlaGln 710
 Db 545 GCCTTCATCAGTCACTGCCTAATAATAATAGCACTAAAGTAGGAGCAAAAGAACTCAG 604
 Qy 711 LeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLys 730
 Db 605 CTCTCTGGTGGCCAGAAACACGATTCGATAGCTGCTGCTGCTGCTGCTGCTGCTCAT 664
 Qy 731 IleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGln 750
 Db 665 ATTTTCTTTGGATGAGCCACGCTCAGCTCTGATACAGAAAGTGAAGAGTTGTCCAA 724
 Qy 751 HisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSer 770
 Db 725 GAAGCCCTGACAAAAGCCAGAGAAGCGCCACCTGATGTTGCTCACCCTGCTGCC 784
 Qy 771 AlaIleGlnAsnAlaAspLeuIleVal-ValLeuHisAsnGlyLysIleLysGluGlnG 790
 Db 785 ACCATCCAGAAATGACACTTAATAGTGTGTTTTCAGAAATGGCAGAGTCAAGGAGCATGG 844
 Qy 790 YThrHisGlnGluLeuLeuArgAsnArgAspIleTyPheLysLeuValAlaGlnSe 810
 Db 845 CAGCATCAGCAGCTGCTGCGACAGAAAGCACTATTTTCAATGCTCAGTGTCCAGGC 904
 Qy 810 r 810
 Db 905 T 905
 RESULT 2
 BM904842
 LOCUS
 DEFINITION BM904842 998 bp mRNA linear EST 12-MAR-2002
 AGENCOURT 6699581 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5557655
 5', mRNA sequence.
 ACCESSION BM904842
 VERSION BM904842.1 GI:19355221
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 998)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs@mail.nih.gov
 Tissue Procurement: ATCC/DCTD/DTP
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM12279 row: f column: 24
 High quality sequence stop: 738.
 Location/Qualifiers
 1. .998
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5557655"
 /clone_lib="NIH_MGC_72"
 /tissue_type="melanotic melanoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 2 kb. Library constructed by Life
 Technologies."
 BASE COUNT 271 a 223 c 248 g 256 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 2.2e-86 Length: 998
 Score: 834.00 Matches: 164

Percent Similarity: 74.12% Conservative: 68
 Best Local Similarity: 52.40% Mismatches: 77
 Query Match: 20.45% Indels: 4
 DB: 14 Gaps: 0

US-09-873-409-2 (1-812) x BM904842 (1-998)

QY 335 AspileAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAla 354
 DB 61 GATGTGAGTTGGTTGATGACCCCTAAAAACACCACTGGAGCATTTGACTACCGGCTGCC 120

QY 355 IleAspileAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsn 374
 DB 121 AATGATGCTCTCAAGTTAAAGGGCGCTATAGGTTCCAGGCTGTGCTGAATTAACCCAGAAT 180

QY 375 AlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyrGluMetThrPhe 394
 DB 181 ATAGCAATCTTGGACAGGAGTAATATATATCTTCATCTATGTTGGCACTAACACTG 240

QY 395 LeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMet 414
 DB 241 TTACTCTTAGCAATTTGACCATCTATGCAATAGCAGGAGTTGTGAATCAAAATGTTG 300

QY 415 ThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGlu 434
 DB 301 TCTGGCAAGCACTGAAAGATAAGAAAGAACTAGAGGCTGTGGGAAGATCGCTACTGAA 360

QY 435 AlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMet 454
 DB 361 GCATAGAAAACCTCCGACCGTGTCTTCTTGACTCAGAGCAGAGTTGTAACATATG 420

QY 455 TyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGly 474
 DB 421 TATGCTCAGAGTTTGCAGGTACCATACAGAAACTCTTTGAGGAAGCACACATCTTTGGA 480

QY 475 SerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPhe 494
 DB 481 ATTACATTTTCTTCCACCGCAATGATGATTTTCTTATGCTGATGATTTTCCGGTTT 540

QY 495 GlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThr 514
 DB 541 GGAGCCTACTTGTGGGCACATAAATCATGAGCTTGGAGATGTTCTGTAGTATTTTCA 600

QY 515 AlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSer 534
 DB 601 GCTGTTCTTGTGTCATGGCGTGGGCAAGTCACTGTTCTCTGCTGCTGACTATGCC 660

QY 535 LysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAsp 554
 DB 661 AAAGCCAAAATATACGACGCGCATCATCATCATGATCATTTGAAAAAACCCTTTGATTGAC 720

QY 555 SerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGlu 574
 DB 721 AGTACAGCAGGAGGAGCCTAATGCCGACACATTTGGAAGAAATGTCACATTTGGTGA 780

QY 575 ValSerPheTyrProCysArgProAspValPheIleLeuArgGlyLeuSer-LeuSe 594
 DB 781 GTTGATTTCAACTATCCACCGACCGACATCCAGTCTTCAGGACTGAGGCCCTTGA 840

QY 594 rIleGluArgGlyLysThrValAlaPheVal-GlySerSerGlyCysGlyLysSerThrS 614
 DB 841 GGTGAAGAGGGGCGAGCCCTGCTGCTGGGGGCGAGTGGCTGTGGGAAGAGACACAG 900

QY 614 er-ValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGly 633
 DB 901 TGGGTCAGCTCTGGAGCGGGTCTCAACCCCTTGGCAAGAAATGCTGCTTGGTGGC 960

QY 634 ValAspAlaLys-GluLeuAsnValGlnTyr 643
 DB 961 AAAAGATNAACCGACTGGATGTTCCATGG 991

RESULT 3
 AK014319

LOCUS AK014319 2676 bp mRNA linear HTC 19-JAN-2002
 DEFINITION Mus musculus 14, 17 days embryo head cDNA, RIKEN full-length
 enriched library, clone:322401P09:ATP-binding cassette, sub-family
 B (MDR/TAP), member 8, full insert sequence.
 ACCESSION AK014319
 VERSION 1 GI:12852089
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (strain:C57BL/6J) 14, 17 days embryo head cDNA to
 mRNA, clone lib:RIKEN full-length enriched mouse cDNA library
 clone:322401P09.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1
 REFERENCE Carninci, P. and Hayashizaki, Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636
 2
 REFERENCE Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE 20499374
 PUBMED 11042159
 3
 REFERENCE Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
 Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M.,
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujikawa, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
 Yoneda, Y., Ishikawa, T., Ogawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 TITLE RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
 MEDLINE 20530913
 PUBMED 11076861
 4
 REFERENCE Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
 Aizawa, K., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
 Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,
 Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
 Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
 Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
 Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Tomita, M.,
 Quackenbush, J., Schriml, L.M., Stauber, F., Suzuki, R., Aono, H.,
 Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
 Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
 Carninci, P., de Bona, M.F., Brownstein, M.J., Bult, C.,
 Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
 Hofmann, M., Hume, D.A., Kaniya, M., Lee, N.H., Lyons, P., Nordone, P.,
 Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Sasaki, H.,
 Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
 Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H.,
 Togo-oka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L.,
 Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohetsuki, S.
 and Hayashizaki, Y.
 TITLE Functional annotation of a full-length mouse cDNA collection
 JOURNAL Nature 409 (6821), 685-690 (2001)
 MEDLINE 21085660
 PUBMED 11217851
 5 (bases 1 to 2676)
 REFERENCE Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,
 Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C.,
 Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T.,
 Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiraoka, T., Hori, F.,
 Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T.,
 Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S.,
 Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K.,
 Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J.,


```
Db 1200 TTCAATGGGGCTCCCTGTGGCTGGACAGACAGCTGAAGGGGAGACCTCATGTCTCTC 1259
Qy 511 ILeValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAla 530
Db 1260 CTGGTGGCTCCAGACAGTACAGAGTCTATGGCC-----AGCTCTCTGTCTCTGTTT 1313
Qy 531 ProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGlyLys 550
Db 1314 GGTCAAGTGTACGTGGCTGAGTGGGGAGCCGAGTCTTCGAATACATACAGCCCTGAGC 1373
Qy 551 ProAsnIleAspSerArgSerGlnGluGlyLysProAspThrCysGluGlyAsnLeu 570
Db 1374 CCGTGTATCCCAATGACCGGGGGCTACTGTGATCCCAACAAAGGACATTCGTGGTCCATC 1433
Qy 571 GluPheArgGluValSerPheThrProCysArgProAspValPheIleLeuAspGly 590
Db 1434 ACCTTCMAAATGTCACTTCACCTACCTCCCTGAGACCTGGCTCAATGTGCTCAAGGAC 1493
Qy 591 LeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGly 610
Db 1494 TTCACCTGAAGCTGCCTCTGCAAGATTGTGGCTCTTGTGGCCAGTCTGGGGAGGA 1553
Qy 611 LysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlnValLeu 630
Db 1554 AAGACCACAGTTCGCTCTGTCTGGAACGCTCTATGACCCCTGAAGCTGGCTCGGTGAGC 1613
Qy 631 PheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGln---IleAla 649
Db 1614 TTGATGGGATGACCTGCGAATCTCAACCTCTCGGCTCGGGCCAGTCAATAGT 1673
Qy 650 IleValProGlnProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGly 669
Db 1674 TTCAATCAGCAGGAGCCAGTCTCTTTGCAACAACCATCATGAGAATATCCGATTTGGG 1733
Qy 670 -----AspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAla 687
Db 1734 AAGCTGATCTTCCGAT-----GAAGAGGTGTACACAGCTGCACGAGAAGCC 1781
Qy 688 AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLys 707
Db 1782 AATGCCACAGTTCATCAGCAGCTTCCCGATGGCTACAGCACTGTGGTTGTGAGCGG 1841
Qy 708 GlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGln 727
Db 1842 GGCACAACCTTGTCTGTGGCCAGACAGCGCTACCCATCGCACGTGCGCTCATCAAG 1901
Qy 728 LysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 747
Db 1902 CAGCCACAGTGTATCTCTGGACAGGCCACCAAGTCCGCTAGATGCAGAATCCGAGAGG 1961
Qy 748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767
Db 1962 GTGGTACAGAGGCCCTGGACCGCGCCAGTGTGGCCGACCGGTGTGGTCATGGCCAC 2021
Qy 768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys 787
Db 2022 CGCTTAGTACTGTCCGTGAGCCCACTCCATCATTTGATGGCCATGGCCAACTGTGT 2081
Qy 788 GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807
Db 2082 GAGGCTGGGACCCACGAAGAACTCTTTAAAAAGCGCGGTCTATTTCAGAGCTTATCCGG 2141
Qy 808 AlaGlnSerVal 811
Db 2142 AGACAAACCTGT 2153
RESULT 4
BG248052 1019 bp mRNA linear EST 13-FEB-2001
LOCUS 60235987F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:4488404 5',
DEFINITION mRNA sequence.
ACCESSION BG248052
```

```
VERSION BG248052.1 GI:12757867
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 1019)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10134 row: n column: 21
High quality sequence stop: 650.
FEATURES
Location/Qualifiers
1..1019
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4488404"
/clone_lib="NCI CGAP Mam1"
/tissue_type="tumor, biopsy sample"
/dev_stage="10 months, virgin"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dr.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"
BASE COUNT 276 a 262 c 283 g 196 t
ORIGIN
Alignment Scores:
Pred. No.: 6.2e-79 Length: 1019
Score: 770.50 Matches: 154
Percent Similarity: 79.17% Conservative: 36
Best Local Similarity: 64.17% Mismatches: 47
Query Match: 18.89% Indels: 3
DB: 12 Gaps: 1
US-09-873-409-2 (1-812) x BG248052 (1-1019)
Qy 567 GluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPhe 586
Db 4 GAAGGAAATGTGCAATTTAGTGGAGTCTGTTCAACTATCCACCGACCCAGCATCCCA 63
Qy 587 IleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySer 606
Db 64 GTCTCTCAGGGGTGAGCTTGTAGGTGAAGAGGGCCAGACGCTGGCCCTGGTGGGGAGC 123
Qy 607 SerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGln 626
Db 124 AGTGGCTGGGGAGAGACACAGTGGTCCAGCTCTCGAGCGCTTCTACGACCCCATGGCT 183
Qy 627 GlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSer 646
Db 184 GGATCATGTCTTAGATGGCAAGAAATAAAGCAACTGAATGTCCAGTGGCTCCGAGCA 243
Qy 647 GlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIle 666
Db 244 CAGCTGGGCATTTGTGCCAAGAGCCCATCTCTTGTACTGCAGCATTCGAGAGAAACATT 303
Qy 667 AlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAla 686
Db 304 GCCTACGGAGACAAACAGCCGGTCTGTCTTATGAGGAGATTTGTAGGGGAGCCCAAGAG 363
Qy 687 AlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeu 706
```



```

Db 364 GCCAACATCACCAGTTTCATCGACTCGTACCTGATAAATACCAACACGAGTAGGAGAC 423
QY 707 LysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaAlaArgAlaLeuLeu 726
Db 424 AAGGGCACTCAGCTGTGGTGGGCAAGCAGCGCATGCCATCGCACCGGCCCTCGTC 483
QY 727 GlnLysProLysLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 746
Db 484 AGACAGCTCACATTTCCTTCTGGACAGCAACATCAGCTCTGGATACAGAAAGTCAA 543
QY 747 LysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThr 766
Db 544 AAGTTTGTCCAGGAAGCGCTGGACAAAGCCAGGAGCGACCTGCATTGTGATCGCT 603
QY 767 HisArgLeuSerAlaLeuGlnAsnAlaAspLeuLeuLeuValValLeuHis-AsnGlyLysI 786
Db 604 CACCGCTGTCCACCATCCCAAGACCGGACTTGATCGGTGGTGGTATTCAAGAACCGGAAG 663
QY 786 LeLys---GluingLysThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 803
Db 664 TCAAGGAGCAGCGGCCACCCCAACAGCAAGTGGTGGCGCAGAGGGCATCTACTTC 719

RESULT 5
BG293345
LOCUS 602390738F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:4502552 5',
DEFINITION mRNA sequence.
ACCESSION BG293345
VERSION BG293345.1 GI:13052943
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 726)
AUTHORS NIH-MGC http://imgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: The Cepko Laboratory
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10371 row: 1 column: 09
High quality sequence stop: 658.
FEATURES
source Location/Qualifiers
1..726
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:4502552"
/clone_lib="NIH_MGC_94"
/tissue_type="retina"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pCMV-SPORT6; Site: 1: NotI;
Site: 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 3.3 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

BASE COUNT 189 a 220 c 190 g 127 t
ORIGIN

Alignment Scores:
Pred. No.: 3,9e-77 Length: 726
Score: 753.00 Matches: 152
Percent Similarity: 84.29% Conservatives: 25
Best Local Similarity: 72.38% Mismatches: 30
Query Match: 18.46% Indels: 3
DB: 12 Gaps: 0

```

```

US-09-873-409-2 (1-812) x BG293345 (1-726)
QY 601 ValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArg 620
Db 2 CTGGCCCTGTGGTGGCAGCAGTGGCTCGGGAAGAGACACAGTGTCTCAGCTGCTCGAGCGC 61
QY 621 LeuTyrAspProValGlnGlnValLeuLeuPheAspGlyValAspAlaLysGluLeuAsn 640
Db 62 TTTATGACCCCATGCTGGATCAGTCTCTTAGATGGTCAAGAAGCAAGAACAATCAAT 121
QY 641 ValGlnTriLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCys 660
Db 122 GTCAGTGGCTCGAGCTCAACTGGGCATTGTCTCCAGGAACCCATTCTCTTTGACTGC 181
QY 661 SerIleAlaGlnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 680
Db 182 AGCATCGCAGAGAACATCGCCTATGGAGAACACAGCGGGTCTGCTCATGATGAGATT 241
QY 681 LysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyr 700
Db 242 GTGAGGCGCAGCAAGGAGGCCACATCCACCTTTCATCGAGACGCTGCCCAAAATAT 301
QY 701 AsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAla 720
Db 302 AACACAAGAGTAGGAGACAAGGGGAGCGCAGCTCTCTGGGGGCCAGAGCAGAGATTGCC 361
QY 721 IleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuLeuLeuLeuLeuLeuLeu 740
Db 362 ATCGCCCGAGCCCTCATCAGACAGCTCGGGTCTCTACTGCTGGATGAAGCAGCTAGCT 421
QY 741 LeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArg 760
Db 422 CTGGATCTCAGAGTGAAGAAGTTGTCCAGGAAGCAGCTGGACAAAGCCAGGAGCGCG 481
QY 761 ThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAla-AspLeuLeuValva 780
Db 482 ACCTGCATTGCGATCGCTCACCGCTGTCCACCATCCAGAACGCGGCACCTTGATCGTGGT 541
QY 780 LLeuHisAsnGlyLysIleLysGluGlnGlyThr-HisGlnGluLeuLeuArgAsnArg 800
Db 542 GATTGAGACGGCAAGGTCAAGGAGCAGCGCACCCACCGAGCTGCTGGCGCAGAAGG 601
QY 800 spIleTyrPhe-LysLeuValAsn 807
Db 602 GCATCTATTCTCAACTGGTCAAC 625

RESULT 6
BF796582
LOCUS 602258463F1 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:4341710 5',
DEFINITION mRNA sequence.
ACCESSION BF796582
VERSION BF796582.1 GI:12101636
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 944)
AUTHORS NIH-MGC http://imgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Louis Staudt, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9955 row: n column: 15

```



```
Db 20 GCAAAGTACTTATGATGTTATTGATAGAAAACCACT-ATTGATTGTTATAGTGAAGAA 78
Qy 560 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 579
Db 79 GGTGAAACATTAAATGATGTTAAAGGTGAAATTAAGATTAAGATTTGTTTCAGATAT 138
Qy 580 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 599
Db 139 CCAACAGACCAGACCAATCTGCTGTGAAGGTATTTTCAATCAAGTAGAACAAAGAAA 198
Qy 600 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuGln 619
Db 199 ACTGTTGCTATTAGTAGGAGCATCAGGATGTGTAATCAACATCAGTTCAGTTGATTGAA 258
Qy 620 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 639
Db 259 AGATTTTATGATCCACACATCGGAGATGATTATTAGTAGACATAATATCAAAAGATTG 318
Qy 640 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 659
Db 319 AATATTCTATTTCTTAAGAGTCAATGGAATGGTAGGACCAAGACCAAGTATTATTGCT 378
Qy 660 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgVal---ValProLeuAsp 678
Db 379 GAAGTGTATTGGATAATATTAGAGAGGAGTACTTAAGAGGTTGAAGTAAGTAAGTAA 438
Qy 679 GluLeLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGlu 698
Db 439 CAAATTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 498
Qy 699 LysTyrAsnThrGlnValGlyLeuLysGlyValGlnLeuSerGlyGlyGlnLysGlnArg 718
Db 499 GGATATAACAAATGTAAGGTATAGAGGTGCACAAATTTTCAGGAGGACAAACAAAGA 558
Qy 719 LeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThr 738
Db 559 ATTGCTATTGTCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 618
Qy 739 SerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThr 758
Db 619 TCAGCAGCTTGTATTCAGAAAGTGAAGATTTGTACAAAGATGCCTTGCACAAAGACAAA 678
Qy 759 GlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIle 778
Db 679 GGAAGAACAAATTTGTAATTCGCATAGATATCACTATTCAAAATGCAGATCAATA 738
Qy 779 ValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsn 798
Db 739 TGTGTTATTATCAGAGGAAGATTCAGAAAGAGGAACACATCAAGAGTATTAGATTG 798
Qy 799 ArgAspIleTyrPheLysLeuValAsnAlaGln 809
Db 799 AAAGGATTTTATTATACACTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 831

RESULT 8
AZ682350
LOCUS ENTKB16TF Entamoeba histolytica Sheared DNA Entamoeba histolytica
DEFINITION genomic, DNA sequence.
ACCESSION AZ682350
VERSION AZ682350.1 GI:11819496
KEYWORDS GSS.
SOURCE Entamoeba histolytica.
ORGANISM Entamoeba histolytica; Entamoebidae; Entamoeba.
REFERENCE 1 (bases 1 to 871)
AUTHORS Loftus,B., Van Aken,S. and Fraser,C.
TITLE Determination of clone end sequences from Entamoeba histolytica
JOURNAL HMI:IMSS sheared DNA library
COMMENT Contact: Brendan J Loftus
Department of Eukaryotic Genomics
```

The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 3543
Email: bjlouf@tigr.org
Clones are derived from the Entamoeba histolytica HMI:IMSS sheared
DNA library

Seq primer: M13-Forward
Class: shotgun
High quality sequence start: 16
High quality sequence stop: 860.

Location/Qualifiers
1. .871
/organism="Entamoeba histolytica"
/strain="HMI:IMSS"
/db_xref="taxon:5759"
/clone_lib="Entamoeba histolytica Sheared DNA"
/notes="Vector: pHOSt; Site 1: Bst I; Constructed at The
Institute for Genomic Research (TIGR), Rockville, MD.
Genomic DNA isolated from broth cultures of E. histolytica
using a method described by Clark and Diamond (Clark,
C.G., and Diamond, D.S. (1993) Entamoeba histolytica: a
method for isolate identification. Exp. Parasitol.
77:450.). The DNA was mechanically sheared to give a
tight size distribution (~2 kb). The v + i method used for
the library construction is described in detail in Smith,
H.O. and Venter, J.C. (Making small insert libraries for
whole genome shotgun sequencing projects. In Genome
Sequencing: A Practical Approach, eds. M. Vaudin and B.
Barelli, Oxford University Press, 1999)."

BASE COUNT 349 a 117 c 174 g 231 t
ORIGIN

Alignment Scores:
Pred. No.: 1.55e-75 Length: 871
Score: 740.50 Matches: 147
Percent Similarity: 70.04% Conservative: 47
Best Local Similarity: 53.07% Mismatches: 82
Query Match: 18.15% Indels: 1
DB: 17 Gaps: 1

US-09-873-409-2 (1-812) x AZ682350 (1-871)

Qy 534 SerLysAlaLysSerGlyAlaAlaHisPheAlaLeuLeuLysLysProAsnIle 553
Db 39 AATAGTCTAAAGTTCGCGCATTTAATGTTTATCAACAATTTGATAGATCCAGATATT 98
Qy 554 AspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArg 573
Db 99 GATTGTCAGTCTATTGGAGGTGAATGTCCAACCTCAGTGAATGGAATATTAGATTTGAG 158
Qy 574 GluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeu 593
Db 159 GATGTTCAATTTGTTTATCCAAAGAGACTGTCTCATCATGTTATTTAAAGGACTTGACCTT 218
Qy 594 SerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThr 613
Db 219 GAAATTAAAGAAAGGACAAACAATTCGATTTAGTTGGAGCATCAGGATGTGGAAATCAACT 278
Qy 614 SerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGly 633
Db 279 ACTATTCAATTAATCCAAAGAAATTAATGATCCAAATGGTGAAGAGTAACATTAGACGA 338
Qy 634 ValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGln 653
Db 339 AAAGATATACGAGAGTTGAATATCAATGGTTAAAGAAATCAATAGGATTAGTTGGACAA 398
Qy 654 GluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArg 673
Db 399 GAACAGATGTTGTTTCAGGAGAACAAATTCGAGAAATTAATTCGTTGGAGCTAAAGAGGA 458
Qy 674 ValValProLeu---AspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPhe 692

Db 459 GCACACCAAGTCAAGAGAGATGATTGAATGTGCTAAATGCGAAATGCAACATGACATCTTC 518
 Qy 693 ileGluGlyLeuProGluGlyTyrAsnThrGlnValGlyLeuGlyGlyAlaGlnLeuSer 712
 Db 519 ATTCTCTAACTCCAGAGGATATGACACAATAATTTGGAGAAAAGGAGCATTTATTATCA 578
 Qy 713 GlyGlyGlnGlyGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnGlyProGlyIleLeu 732
 Db 579 GGAGGACAAACAAAGAAATGCAATTTGACGTCGATTCATTCGAACCCCATCTATTCTT 638
 Qy 733 LeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluGlyValValGlnHisAla 752
 Db 639 CTCTCTGATGAAGCTACATCAGACCTTGATACACAAGTGAAAAGATTGTACAAAGACA 698
 Qy 753 LeuAspGlyAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIle 772
 Db 699 CTTGAAAAGAGCTCTAAAGGAAGACAACAATAATTTGTAGCACATAGACTTAACAACCTGT 758
 Qy 773 GlnAsnAlaAspLeuIleValValLeuHisAsnGlyGlyIleGlyGlnGlnThrHis 792
 Db 759 AGAAATGCGAGATAAAATATGTGTATTTCATCAAGGAGAAATTTTGAACAAAGGAAACAT 818
 Qy 793 GlnGluLeuLeuArgAsnArgAspIleTyrPheLeuLeuValAsnAlaGln 809
 Db 819 CAAGAATTATGCAATTGTAAGGAACAATATTATGGATTAGTCAAAAGACAA 869

RESULT 9
 AZ670821/c
 LOCUS ENTJUN69TF Entamoeba histolytica Sheared DNA Entamoeba histolytica
 DEFINITION Genomic, DNA sequence.
 ACCESSION AZ670821
 VERSION AZ670821.1 GI:11807967
 KEYWORDS GSS.
 SOURCE Entamoeba histolytica.
 ORGANISM Entamoeba histolytica.
 Eukaryota; Entamoebidae; Entamoeba.
 1 (bases 1 to 932)
 Lofthus, B., Van Aken, S. and Fraser, C.
 Determination of clone end sequences from Entamoeba histolytica
 HM1-IMSS sheared DNA library
 Unpublished (2000)
 Contact: Brendan J Loftus
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0208
 Fax: 301 838 3543
 Email: b.loftus@tigr.org
 DNA library
 Clones are derived from the Entamoeba histolytica HM1-IMSS sheared
 DNA library
 Seq primer: M13-Forward
 Clases: shotgun
 High quality sequence start: 46
 High quality sequence stop: 867.
 Location/Qualifiers
 1..932
 /organism="Entamoeba histolytica"
 /strain="HM1-IMSS"
 /db_xref="taxon:5759"
 /clone_lib="Entamoeba histolytica Sheared DNA"
 /note="Vector: pHOS1; Site 1: Bst I; Constructed at The
 Institute for Genomic Research (TIGR), Rockville, MD.
 Genomic DNA isolated from broth cultures of E. histolytica
 using a method described by Clark and Diamond (Clark, a
 C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
 method for isolate identification. Exp. Parasitol.
 77:450.). The DNA was mechanically sheared to give a
 tight size distribution (~2 kb). The v + i method used for
 the library construction is described in detail in Smith,
 H.O. and Venter, J.C. (Making small insert libraries for
 whole genome shotgun sequencing projects. In Genome
 Sequencing: A Practical Approach, eds. M. Vaudin and B.

BASE COUNT 258 a 182 c 125 g 367 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 8,76e-75 Length: 932
 Score: 734.50 Matches: 150
 Percent Similarity: 71.11% Conservative: 42
 Best Local Similarity: 55.56% Mismatches: 77
 Query Match: 18.01% Indels: 2
 DB: 17 Gaps: 1
 US-09-873-409-2 (1-812) x AZ670821 (1-932)
 Qy 533 TyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGlyLysProAsn 552
 Db 835 TTTTGCTACAGCTAAAGCTTCTGCATATAGAAATTTATCAACAATTCATAGATCCAGAT 776
 Qy 553 IleAspSerArgSerGlnGluGlyLysProAspThrCysGluGlyAsnLeuGluPhe 572
 Db 775 ATTGATTGTAGATCTACAGCTGGTGAATGTCCAACCTGAGTGAATGGAATATTACATTA 716
 Qy 573 ArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSer 592
 Db 715 GAAGATGTTCAATTTAGATATCCACAAGACCACAAACAATTCCTGGTGACCTTGAT 656
 Qy 593 LeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSer 612
 Db 655 CTTGAAATTTAAGAAAGGACAAACAGTTGCTATTAGTAGGAGCATCAGGATGTGTAATCA 596
 Qy 613 ThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAsp 632
 Db 595 ACTACTATTCAATTAGTCCAAAGAAATTTATGATCCAGTTGGTGGGTCAATTAATAGAT 536
 Qy 633 GlyValAspAlaLysGluLeuAsnValGlnTyrLeuArgSerGlnIleAlaIleValPro 652
 Db 535 GGAAAGATTTAAGAGATTTAAATATCAATGTTAAGAAATCAATAGATTAGTTGGA 476
 Qy 653 GlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSer 672
 Db 475 CAAGAACCTATTTTGTTGCTATGCTACTATTAGAGAAATATTATTGCTGGAGCTAGAT 416
 Qy 673 ArgValValProLeu---AspGluIleLysGluAlaAlaAsnAlaAlaIleHisSer 691
 Db 415 GGAGAAACCACTCAAGAGAGAGATGATTGAATGTGCTAAATATGCAATGCAATGAA 356
 Qy 692 PheIleGluGlyLeuProGluGlyTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeu 711
 Db 355 TTTATTCTCATCTTCCAGAGGATATGATACAATGTTAGGAGAAAAGAGCTGCATTA 296
 Qy 712 SerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIle 731
 Db 295 TCAGGAGGACAAACAAAGAAATGCTATTGCACGTGCTATTGATTAGAAACCTACAAT 236
 Qy 732 LeuLeuLeuAspGluAlaThrSerAlaLeuAspAspSerGluLysValValGlnHis 751
 Db 235 TTATTACTTGTAGTGAAGCTACATCAGCACTTGATACACAAGTGAAAAGATTGTACACAA 176
 Qy 752 AlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAla 771
 Db 175 GCATTTGAAAAGAGCTAGTCAAGAGAAACAACATTTGTTGTAGCACATAGATTAACT 116
 Qy 772 IleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThr 791
 Db 115 GTTAGAAATGCAAGTAGAATTTCTGTATTCCATCAAGGAGAAATTTAT-GAACAAGGACA 57
 Qy 792 HisGlnGluLeuLeuArgAsnArgAspIle 801
 Db 56 CATCAAGATTAAAGGATTTAAAGGATTTA 27
 RESULT 10
 AZ683753
 LOCUS

DEFINITION ENT196TF Entamoeba histolytica Sheared DNA Entamoeba histolytica genomic, DNA sequence.

ACCESSION AZ683753

VERSION AZ683753.1 GI:11820899

KEYWORDS GSS.

SOURCE Entamoeba histolytica.

ORGANISM Entamoeba histolytica.

REFERENCE 1 (bases 1 to 947)

AUTHORS Loftus,B., Van Aken,S. and Fraser,C.

TITLE Determination of clone end sequences from Entamoeba histolytica HM1:IMSS sheared DNA library

JOURNAL Unpublished (2000)

COMMENT Contact: Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 3543
Email: bjlloftus@tigr.org
Clones are derived from the Entamoeba histolytica HM1:IMSS sheared DNA library
Seq primer: M13-Forward
Class: shotgun
High quality sequence start: 40
High quality sequence stop: 778.

FEATURES
source
1..947
/organism="Entamoeba histolytica"
/strain="HM1:IMSS"
/db_xref="taxon:5759"
/clone_lib="Entamoeba histolytica Sheared DNA"
/notes="Vector: pHOS1; Site 1: Bst I; Constructed at The Institute for Genomic Research (TIGR), Rockville, MD
Genomic DNA isolated from broth cultures of E. histolytica using a method described by Clark and Diamond (Clark, C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a method for isolate identification. Exp. Parasitol. 77:450.). The DNA was mechanically sheared to give a tight size distribution (~2 kb). The v + i method used for the library construction is described in detail in Smith, H.O. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B. Barrell, Oxford University Press, 1999)."

BASE COUNT 379 a 121 c 178 g 269 t

ORIGIN

Alignment Scores:
Pred. No.: 4.54e-74 Length: 947
Score: 728.50 Matches: 149
Percent Similarity: 69.82% Conservative: 43
Best Local Similarity: 54.18% Mismatches: 81
Query Match: 17.86% Indels: 2
DB: 17 Gaps: 1

US-09-873-409-2 (1-812) x AZ683753 (1-947)

Qy 536 AlaLysSerGlyAlaAlaHisLeuPheAlaLeuGluLysProAsnLeuSer 555

Db 69 GCAAAAGCATCAGCCTATAAATATTTACACAAATTCAGATATCCAGATATTGTT 128

Qy 556 ArgSerGlnGluLysLeuProAspThrCysGluGlyAsnLeuGluPheArgGluVal 575

Db 129 CAGTCTATTGGAGTGAATGCCAATGAGTGAATATTTAGATTGATGTT 188

Qy 576 SerPheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerile 595

Db 189 CAATTGTTTATCCACAGACCATCTCATCATGTTATTAAGGAATGACCTTGAATT 248

Qy 596 GluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerVal 615

Db 249 AAGAAAGGAGAAATAATTCATTAGTTGGAGCATCAGGATGTAGAAAGTCAACTACTATT 308

Qy 616 GlnLeuLeuGlnArgLeuTyrAspProValGlnGlnGlyGlnValLeuPheAspGlyValAsp 635

Db 309 CAATTAAATCAAAGAAATATGAACCAATGGTGGAGAGTAGTAACATTAGATGGAAAGAC 368

Qy 636 AlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluPro 655

Db 369 ATAAGAGACTTGAATATCAATGGTTTAAAGAAATCAATAGGAATAGTTGGACAAAGACCA 428

Qy 656 ValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGly---AspAsnSerArgVal 674

Db 429 GTATTATTTTCAGGAACAATTTGGAGAAATATCATCTCTGGAGCTTAAGAGAGAAACA 488

Qy 675 ValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGlu 694

Db 489 CTAAGTAAAGAGAGATGATTGAATGTACTAAATGGCAATGCACATGACTTATTCTTCT 548

Qy 695 GlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGly 714

Db 549 AAACCTCCAGAGAGATATGACACAAATAATTGGAGAAAAAGGAGCATTTATTCAGAGGA 608

Qy 715 GlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeu 734

Db 609 CAAAAACAAGAAATTCGAATTCACGTGCTGATTGATTCGAAACCATCTATTCTTCTCT 688

Qy 735 AspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAsp 754

Db 669 GATGAAGCTACATCAGCACTTGATACACAAAGTGAAGAAAGATTGTACAAGAACGACTTGAC 728

Qy 755 LysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsn 774

Db 729 AAAGCATCAAAAGGAAGAACAAATTTATTGTAGGACATAGACTATCGACTATTCAAAAT 788

Qy 775 AlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlu 794

Db 789 GCAGATCAATATGTTTATTATTCAGAGGAAAAAATAGTAGAACAGACACATCAAGAA 848

Qy 795 -LeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAla 808

Db 849 TTTATTGGATTTGAAAAGATTTTATTATACACTTGCTTATGCA 891

RESULT 11

LOCUS AZ540627

DEFINITION ENTQ18TF Entamoeba histolytica Sheared DNA Entamoeba histolytica genomic, DNA sequence.

ACCESSION AZ540627

VERSION AZ540627.1 GI:11147603

KEYWORDS GSS.

SOURCE Entamoeba histolytica.

ORGANISM Entamoeba histolytica.

REFERENCE 1 (bases 1 to 886)

AUTHORS Loftus,B., Van Aken,S. and Fraser,C.

TITLE Determination of clone end sequences from Entamoeba histolytica HM1:IMSS sheared DNA library

JOURNAL Unpublished (2000)

COMMENT Contact: Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 3543
Email: bjlloftus@tigr.org
Clones are derived from the Entamoeba histolytica HM1:IMSS sheared DNA library
Seq primer: M13-Forward
Class: shotgun
High quality sequence start: 23
High quality sequence stop: 856.

FEATURES
source
1..886
/organism="Entamoeba histolytica"

H.O. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In *Genome Sequencing: A Practical Approach*, eds. M. Vaudin and B. Barrell, Oxford University Press, 1999).

genomic, DNA sequence.

ACCESSION Z6687805

VERSION Z6687805.1 GI:11824951

KEYWORDS GSS.

SOURCE Entamoeba histolytica.

ORGANISM Entamoeba histolytica.
Eukaryota; Entamoebidae; Entamoeba.

REFERENCE 1 (bases 1 to 880)
Loftus, S., Van Aken, S. and Fraser, C.
Determination of clone end sequences from Entamoeba histolytica

JOURNAL Unpublished (2000)

Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA

Fax: 301 838 3543
Email: bjloftus@tigr.org
Clones are derived from the Entamoeba histolytica HM1:IMSS strain

Seq primer: M13-Forward
Class: shotgun

High quality sequence stop: 858.

FEATURES	Location/Qualifiers
source	1..880

```

/organism="Laccaseobd histolytica"
/strain="HML:IMSS"
/db_xref="taxon:5759"

```

/note=Vector: pHOS1; Site_1: Bst I; Constructed at Institute for Genomic Research (TIGR), Rockville, MD. Genomic DNA isolated from broth cultures of *E. histolytica*.

using a method described by Clark and Diamond (1970), C.G., and Diamond, L.S. (1993) Entamoeba histolytica method for isolate identification. Exp Parasitol.

tight size distribution (~2 kb). The *v + i* method uses the library construction described in detail in Scazzari et al. (1990) and Venter et al. (1989). (Making small insert libraries.)

whole genome shotgun sequencing projects. In *Genome Sequencing: A Practical Approach*, eds. M. Vaudin and B. Borel, Oxford University Press, 1999).

ORIGIN

Alignment Scores:

Prod. NO.:	1.54E-73	Length:	880
Score:	723.50	Matches:	151
Percent Similarity:	69.93%	Conservative:	42

Query Match:	17.74%	Indels:	2
DB:	17	Gaps:	1

US-09-873-409-2 (1-812) X A268/805 (1-880)

Db 423 ATGTATGCTCAGAGTTTGCGAGGTACCATACAGAACTCTTTGAGGAAAGACACACATCTTT 482
 Qy 474 GlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArg 493
 Db 483 GGAATTACATTTCTCTTACCAGGCAATGATGATATTTCTCTATGCTGATGTTTCCGG 542
 Qy 494 PheGlyAlaTyrIleuLeuGlnAlaGlyArgMetThrProGluGlyMetPheIleValPhe 513
 Db 543 TTTGGAGCCTACTTGTGGGCACATAAACTCATGAGCTTTCAGGATGTTCTGTTAGTATT 602
 Qy 514 ThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrIleuValLeuAlaProGluTyr 533
 Db 603 TCAGCTGTGTTCTTGGTGCATGCGCGTGGGCAAGTCAGTTTCATTTGCTCTGCTGAT 662
 Qy 534 Ser-LysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIle 553
 Db 663 GCCCAAGCCAAATATACAGCCACATCATCATGATCATTTGAAAACCCCTTTGAT 722
 Qy 553 eAepSerArgSer-GlnGluGlyLysPheProAspThrCysGluGlyAsnLeuGlu---P 572
 Db 723 TGACAGCTACAGCAGCCGAGGCGCTAATGCCGAACACATTTGGAAGGAATGTCCACATTT 782
 Qy 572 heArgGluValSerPhePheTyrPro 580
 Db 783 GGTGAAAGTTGATTTCAACTATATCCC 808
 RESULT 14
 LOCUS AZ541090 897 bp DNA linear GSS 14-NOV-2000
 DEFINITION ENTDS67TR Entamoeba histolytica Sheared DNA Entamoeba histolytica genomic, DNA sequence.
 ACCESSION AZ541090
 VERSION AZ541090.1 GI:11148493
 KEYWORDS GSS.
 SOURCE Entamoeba histolytica.
 ORGANISM Entamoeba histolytica
 REFERENCE 1 (bases 1 to 897)
 AUTHORS Lofcus, B., Van Aken, S. and Fraser, C.
 TITLE Determination of clone end sequences from Entamoeba histolytica
 JOURNAL HM1:IMSS sheared DNA library
 COMMENT Unpublished (2000)
 Contact: Brendan J Loftus
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0208
 Fax: 301 838 3543
 Email: bjloftus@tigr.org
 Clones are derived from the Entamoeba histolytica HM1:IMSS sheared DNA library
 Seq primer: M13-Reverse
 Class: shotgun
 High quality sequence start: 81
 High quality sequence stop: 890.
 Location/Qualifiers
 1. 897
 /organism="Entamoeba histolytica"
 /strain="HM1:IMSS"
 /db_xref="taxon:5759"
 /clone_lib="Entamoeba histolytica Sheared DNA"
 /note="Vector: pHOS1; Site 1; Bst I; Constructed at The Institute for Genomic Research (TIGR), Rockville, MD.
 Genomic DNA isolated from broth cultures of E. histolytica using a method described by Clark and Diamond (Clark, C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a method for isolate identification. Exp. Parasitol. 77:450.). The DNA was mechanically sheared to give a tight size distribution (~2 kb). The v + i method used for the library construction is described in detail in Smith, H.O. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.

BASE COUNT 335 a 110 c 177 g 275 t
 ORIGIN
 Alignment Scores: 3.58e-73 Length: 897
 Pred. No.: 720.50 Matches: 143
 Score: 74.59% Conservative: 39
 Percent Similarity: 58.61% Mismatches: 61
 Best Local Similarity: 17.66% Indels: 1
 Query Match: 17 Gaps: 1
 DB: 1
 US-09-873-409-2 (1-812) x AZ541090 (1-897)
 Qy 567 GluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPhe 586
 Db 94 AAAGGTGAATTTGAATTTAAAGACATTTGTTTCAGATATCCAAAGACACAGCAATTTCT 153
 Qy 587 IleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySer 606
 Db 154 GTCTTGAAGAGTATTTCATTCATCAAGTAGAACCAAGGAAAACTGTTGTCATTAGTAGGCA 213
 Qy 607 SerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGln 626
 Db 214 TCAGGATGGTGAATCAACATCAGTTGATTTGAAAGATTTTATGATCCACACAT 273
 Qy 627 GlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSer 646
 Db 274 GAGATGATTTATTAGTAGGACATATATCAAGATTTGAAATATTCATTTCTTTAGAGT 333
 Qy 647 GlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIle 666
 Db 334 CAATTTGGAATGTTAGGACCAAGACCCAGTATTATTGCTGAAAGTGTATTGGATAATATT 393
 Qy 667 AlaTyrGlyAspAsnSerArgVal---ValProLeuAspGluIleLysGluAlaAsn 685
 Db 394 AGAAGAGGAGTACCTAAAGGAGTTGAAGTAAGTAATGAACAAATTTATGCTGCTGCTAAA 453
 Qy 686 AlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGly 705
 Db 454 ATGGCAATGTCATGACTTTTATTTTCAGCAATGCCAGAGGATATTAACAATGTTAGGT 513
 Qy 706 LeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeu 725
 Db 514 GATAGAGGTGCACAAATTTTCAGGAGGACAAAAACAAGAAATGCTATTGACGCTGATTG 573
 Qy 726 LeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSer 745
 Db 574 ATTAGAAATCCAAAAGTGTATTACTCGATGAAGCTACATCAGCACTTGATTCAGAAAGT 633
 Qy 746 GluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValVal 765
 Db 634 GAAAGATTGTCAAGATGCACTTTGACAAGCGCAAGGAGGAGGAGGAGGAGGAGGAGGAG 693
 Qy 766 ThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisGlyLys 785
 Db 694 GCACATAGATTATCAACTATTCAAAATGTCAGATCAAAATGTTGTTATTATGAGAGGAGA 753
 Qy 786 IleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeu 805
 Db 754 ATTGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 813
 Qy 806 ValAsnAlaGln 809
 Db 814 GCTATGCAACAA 825
 RESULT 15
 LOCUS BH155700 913 bp DNA linear GSS 24-SEP-2001
 DEFINITION ENTDS67TR Entamoeba histolytica Sheared DNA Entamoeba histolytica genomic, DNA sequence.
 ACCESSION BH155700
 VERSION BH155700.1 GI:15727822

KEYWORDS
SOURCE Entamoeba histolytica.
ORGANISM Entamoeba histolytica
Eukaryota; Entamoebidae; Entamoeba.
REFERENCE 1 (bases 1 to 913)
AUTHORS Loftus, B., Wang, Z., Van Aken, S. and Fraser, C.
TITLE Determination of clone end sequences from Entamoeba histolytica
HMI:IMSS sheared DNA library (2001)
JOURNAL Unpublished (2001)
COMMENT Contact: Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 3543
Email: bjloftus@tigr.org
Clones are derived from the Entamoeba histolytica HMI:IMSS sheared
DNA library
Seq primer: M13-Reverse
Class: shotgun
High quality sequence start: 7
High quality sequence stop: 805.
Location/Qualifiers
1..913
/organism="Entamoeba histolytica"
/strain="HMI:IMSS"
/db_xref="taxon:5759"
/clone_lib="Entamoeba histolytica Sheared DNA"
/notes="Vector: pHOsi; Site 1: Bst I; Constructed at The
Institute for Genomic Research (TIGR), Rockville, MD.
Genomic DNA isolated from broth cultures of E. histolytica
using a method described by Clark and Diamond (Clark,
C.G. and Diamond, L.S. (1993) Entamoeba histolytica: a
method for isolate identification. Exp. Parasitol.
77:450.). The DNA was mechanically sheared to give a
tight size distribution (~2 Kb). The v + i method used for
the library construction is described in detail in Smith,
H.O. and Venter, J.C. (Making small insert libraries for
whole genome shotgun sequencing projects. In Genome
Sequencing: A Practical Approach, eds. M. Vaudin and B.
Barell, Oxford University Press, 1999)."

BASE COUNT 359 a 118 c 185 g 251 t
ORIGIN

Alignment Scores:
Pred. No.: 4.17e-72 Length: 913
Score: 711.50 Matches: 146
Percent Similarity: 69.37% Conservative: 42
Best Local Similarity: 53.87% Mismatches: 82
Query Match: 17.44% Indels: 2
DB: 17 Gaps: 1

US-09-873-409-2 (1-812) x BH155700 (1-913)

QY 533 TyrSerLysAlaLysSerGlyValAlaHisLeuPheAlaLeuLeuGluLysLysProAsn 552
Db 29 TTTGCTACAGCTAAAGCTTCTGCATATAGATTATCAACAATTATAGATCCAGAT 88
QY 553 IleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPhe 572
Db 89 ATTGATTGTAGATCTACAGCTGGTGAATGTCCAACTGAGTGTAAATGGAATATTACATTA 148
QY 573 ArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSer 592
Db 149 GAAGATGTTCAATTAGATATATCCAAACAGACCAACTAAACAAATCTTGGTGGACTTGAT 208
QY 593 LeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSer 612
Db 209 CTTGAATTAAGAAAGGACAAACAGTTCATTAGTAGGAGCATCAGGATGTGGTAAATCA 268
QY 613 ThrSerValGlnLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAsp 632
Db 269 ACTACTATTCAATTAGTCCAAAGAAATTTATGATCCAGTTGGTGGTCAAGTCAAAATTAGAT 328

QY 633 GlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValPro 652
Db 329 CGAAAAGATTAAAGAGATTAAATATCAAAATGTTAAGAAATCAAAATAGGATTAGTTGA 388
QY 653 GlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSer 672
Db 389 CAAGAACCTTATTTGTCATGTACTATTAGAGAAATATTATTATGCTTGAGCTAGAGAT 448
QY 673 ArgValValProLeu---AspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSer 691
Db 449 CGAAACACCACTCAAGACAGAGATGATTGAATGCTTAAATGCAAAATGCACATGAA 508
QY 692 PheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeu 711
Db 509 TTTATTTCTCATCTTCCAGAGGATATGATACAAATGTTAGGAGAAAAGAGGTGCATTA 568
QY 712 SerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuGlnLysProLysIle 731
Db 569 TCAGGAGGACAAACAAAGAAATTTGCTATTCACGTCGCTATTGATTAGAAACCTCAAT 628
QY 732 LeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHis 751
Db 629 TTATTACTTGTAGAGCTACATCAGCACTTGATACACAAGTGAAAAGATTGTACACAA 688
QY 752 AlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAla 771
Db 689 GCACCTTGAAGAAAGCTAGTCAAGGAGAACAAACAATTTGTTGTAGCACATAGATTAACT 748
QY 772 IleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThr 791
Db 749 GTTAGAAATGCAAGTAGAATTTGTTATTCCTCATCAAGGAGAAATTTAT-TTGACAGGAACA 807
QY 792 HisGlnGluLeuLeuArgAsnArgAspIleTyr 802
Db 808 CATTCAAGATTAAATGATTTGAAGGGGACATAT 840

Search completed: March 31, 2003, 13:49:57
Job time : 5612.21 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 31, 2003, 02:24:20 ; Search time 118.913 Seconds
(without alignments)
2094.145 Million cell updates/sec

Title: US-09-873-409-2
Perfect score: 4079
Sequence: 1 MVDENDRALNVRHDIHG.....QELLNRDIYFKLVNAQSVQ 812

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues
Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPTO.spool/US09873409/runat_27032003_115422_19259/app_query.fasta_1.7544
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09873409 @CGN 1 1.578 @runat_27032003_115422_19259 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LOGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DSEXT=7

Database : Issued Patents NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PTCUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2436	59.7	4264	2 US-08-784-649A-1	Sequence 1, Appli
2	2436	59.7	4264	2 US-08-784-649A-5	Sequence 5, Appli
3	2436	59.7	4646	1 US-08-181-471-2	Sequence 2, Appli
4	2436	59.7	4669	6 5206352-3	Patent No. 5206352
5	2436	59.7	6505	2 US-08-793-610-5	Sequence 5, Appli
6	2436	59.7	9318	2 US-08-793-610-6	Sequence 6, Appli
7	2413.5	59.2	4233	3 US-09-120-513-1	Sequence 1, Appli
8	2413.5	59.2	4233	4 US-09-450-105-1	Sequence 1, Appli
9	2412	59.1	4669	2 US-08-752-447-1	Sequence 1, Appli
10	2412	59.1	4669	4 US-09-316-167-1	Sequence 1, Appli
11	2397	58.8	4669	2 US-08-583-276-18	Sequence 18, Appli
12	2280	55.9	2726	1 US-08-461-823-1	Sequence 1, Appli

13	1690.5	41.4	4002	2 US-08-996-545-1	Sequence 1, Appli
14	1690.5	41.4	4002	2 US-08-996-545-3	Sequence 3, Appli
15	1690.5	41.4	4002	4 US-09-328-320-1	Sequence 1, Appli
16	1690.5	41.4	4002	4 US-09-328-320-3	Sequence 3, Appli
17	1661	40.7	4047	2 US-08-612-734B-1	Sequence 1, Appli
18	1618	39.7	4800	2 US-08-612-734B-3	Sequence 3, Appli
19	1570.5	38.5	3924	1 US-08-395-246C-1	Sequence 1, Appli
20	1474	36.1	4224	1 US-08-612-521-1	Sequence 1, Appli
21	1255.5	30.8	3909	1 US-08-232-537-1	Sequence 1, Appli
22	1223.5	30.0	3924	2 US-08-996-644-3	Sequence 3, Appli
23	1223.5	30.0	3924	3 US-09-352-552-3	Sequence 3, Appli
24	1223.5	30.0	3927	2 US-08-996-644-1	Sequence 1, Appli
25	1223.5	30.0	3927	3 US-09-352-552-1	Sequence 1, Appli
26	1172	28.7	6143	1 US-08-612-521-3	Sequence 3, Appli
27	1116.5	27.4	3792	4 US-09-351-224E-10	Sequence 10, Appli
28	1032.5	25.3	3999	4 US-09-351-224E-9	Sequence 9, Appli
29	915	22.4	13188	4 US-08-961-527-70	Sequence 70, Appli
30	824	20.2	6492	4 US-08-961-527-188	Sequence 188, Appli
31	793	19.4	7760	4 US-08-961-527-63	Sequence 63, Appli
32	776	19.0	7186	4 US-08-961-527-39	Sequence 39, Appli
33	763	18.7	4403765	4 US-09-103-840A-2	Sequence 2, Appli
34	763	18.7	4411529	4 US-09-103-840A-1	Sequence 1, Appli
35	678	16.6	2376	1 US-08-394-880B-1	Sequence 1, Appli
36	644	15.8	2244	4 US-09-061-764A-18	Sequence 18, Appli
37	623.5	15.3	2061	4 US-09-061-764A-17	Sequence 17, Appli
38	608.5	14.9	4781	2 US-09-001-273-1	Sequence 1, Appli
39	608.5	14.9	4781	4 US-08-843-459A-1	Sequence 1, Appli
40	608.5	14.9	4847	3 US-09-061-400-1	Sequence 1, Appli
41	608.5	14.9	5120	3 US-08-772-270A-6	Sequence 6, Appli
42	608.5	14.9	8370	2 US-08-488-706-1	Sequence 1, Appli
43	601	14.7	1749	4 US-09-134-001C-1893	Sequence 1893, Appli
44	600.5	14.7	1743	4 US-09-134-001C-2774	Sequence 2774, Appli
45	594.5	14.6	7721	3 US-08-772-270A-14	Sequence 14, Appli

ALIGNMENTS

RESULT 1
US-08-784-649A-1
; Sequence 1, Application US/08784649A
; Patent No. 5830697
; GENERAL INFORMATION:
; APPLICANT: Sikic, Branimir I
; TITLE OF INVENTION: P-GLYCOPROTEIN MUTANT RESISTANT TO
; TITLE OF INVENTION: CYCLOSPORIN MODULATION
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 2200 Sand Hill Road
; CITY: Menlo Park
; STATE: CA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/784,649A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: Reg. No. 5830697 36,677
; REFERENCE/DOCKET NUMBER: 06037/007001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-322-5070
; TELEFAX: 415-854-0875
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4264 base pairs

; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 US-08-784-649A-1

Alignment Scores:

Pred. No.: 6,88e-275 Length: 4264
 Score: 2436.00 Matches: 466
 Percent Similarity: 76.79% Conservative: 166
 Best Local Similarity: 56.62% Mismatches: 177
 Query Match: 59.72% Indels: 14
 DB: 2 Gaps: 4

US-09-873-409-2 (1-812) x US-08-784-649A-1 (1-4264)

QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
 DB 1495 GTTGATGGACAGGATATTAGGACCATTAATGTAAGGTTTCTACGGGAATCATTTGGTGTG 1554
 QY 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
 DB 1555 GTGAGTCAGAACTGATGTTGTTGTCACACAGATAGCTGAAGAACATTCGCTATGGCCGT 1614
 QY 42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
 DB 1615 GAAATGTCACCATGATGAGTTGAGAAAGCTGTCAAGGAAGCAATGCCTATGACTTT 1674
 QY 62 IleMetGluPheProAsnLysPheAsnThrIleuValGlyGluLysGlyAlaGlnMetSer 81
 DB 1675 ATCATGAACTGCCTCAATAATTTGACACCCCTGGTTGGAGAGAGGGGCCAGTTGAGT 1734
 QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
 DB 1735 GTGGGAGAGCAGAGATCGCATTTGACGTCCTGTTGCGAACCCCAAGATCCTC 1794
 QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
 DB 1795 CTGCTGATGAGGCCACGTGAGCCTTGGACACAGAAAGCAGCGTTCAGTGGCT 1854
 QY 122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
 DB 1855 CTGATAAGGCCAGAAAGGTGCGACCACTTGTGATAGCTCATCGTTTGTCTACAGTT 1914
 QY 142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
 DB 1915 CGTAATGTCACGTCATCGCTGGTTTCGATGATGAGTCATCTGGGAAGGAATCAT 1974
 QY 162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln----- 178
 DB 1975 GATGAACCTCATGAAGAGAAAGGACATTTACTCAAACCTTGTCAATGCAGACAGCAGGA 2034
 QY 179 -----AspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu 195
 DB 2035 AATGAAGTTGAAATTAGAAATGACGCTGATGAATCCAAAAGTGAAATGATGCTGGAA 2094
 QY 196 ArgLysThrAsnSerLeuProLeuHisSerVal-----LysSerIle 209
 DB 2095 ATGCTTCAATGATTCAGATCCAGTCTTAATAGAAAAGATCAACTCTGAGAGTGTG 2154
 QY 210 Lys---SerAspPheIleAspLysAlaGluSerThrGlnSerLysGluIleSerLeu 228
 DB 2155 CGTGGATCAACAGCCCAAGACAGAAAGCTTAGTACCAAGAGGCTCTGGATGAAGTATA 2214
 QY 229 ProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValVal 248
 DB 2215 CCTCCAGTTTCTTTGGAGGATTTAAGCTAAATTTAACTGAATGGCCTTATTTGTT 2274
 QY 249 LeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePhe 268
 DB 2275 GTTGGGTATTTTGTGTCATTTAATAATGGAGCCCTGCAACAGCATTTGCAATAATTT 2334
 QY 269 AlalysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeuLysHisAspAla 287

DB 2335 TCAAGAGATTATAGGGGTTTTTACAAGAAATTGATGATCCTGAAACAAACACAGAAATAGT 2394
 QY 288 GluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMet 307
 DB 2395 AACTTGTTCACATATTGTTTCTAGCCCTGGAAATATTTCITTTTATACATTTCCCTT 2454
 QY 308 GlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla 327
 DB 2455 CAGGGTTTCACATTTGGCAAAGCTGGAGAGATCCTCACCNAAGGGCTCCGATACATGGTT 2514
 QY 328 PheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGly 347
 DB 2515 TTCGATCCATGCTCAGACAGAGATGTGTTGTTGATGACCTTAAACACACCACTGGA 2574
 QY 348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg 367
 DB 2575 GCATTGACTACACAGGCTGCCAATGATGTGCTCAAGTTAAAGGGCTATAGTTCCAGG 2634
 QY 368 IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387
 DB 2635 CTGCTGTAATTAACCAAGATATAGCAATCTTGGACAGGAATATATATCTTCATC 2694
 QY 388 TyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGly 407
 DB 2695 TATGGTTGGCAACTAACACTGTTACTCTTAGCAATTTGTACCCATCATTTGCAATAGCAGGA 2754
 QY 408 MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHis 427
 DB 2755 GTTGTGGAATGAAATGTTGCTCGCAAGCACTGAAAGATAGAAAGAACATAGAGGT 2814
 QY 428 AlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArg 447
 DB 2815 GCTGGGAAGATCGCTACTGAAGCAATAGAAACTTCGGAACCGTTGTTCTTTGACTCAG 2874
 QY 448 GluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSer 467
 DB 2875 GAGCAGAAGTTTGAAACATATGATGCTCAGAGTTTCAGGTACCATACAGAAACTCTTTG 2934
 QY 468 LysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487
 DB 2935 AGAAAGCACACATCTTTTGGAAATACATTTCTCCACAGCAATGATGATTTTTC 2994
 QY 488 TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507
 DB 2995 TATGCTGGATGTTTCGGTTTGGAGCTACTTGTGGCACATAACTCATGACCTTTGAG 3054
 QY 508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeu 527
 DB 3055 GATGTTCTGTAGTATTTTTCAGCTGTGTTGTTGGTGCCATGGCGCGGCAAGTCAGT 3114
 QY 528 ValLeuAlaProGluTyrSerLysAlaLysSerGlyValAlaAlaHisLeuPheAlaLeuLeu 547
 DB 3115 TCATTTGCTCCTGACTATGCCAAAGCCAAATATACAGCAGCCACATCATCATGATCAT 3174
 QY 548 GluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGlu 567
 DB 3175 GAAAAACCCCTTTGATTGACAGCTACAGCAGCGAAGGCTTAATGCCGAACACATTTGGA 3234
 QY 568 GlyAsnLeuGluPheArgGluValSerPheTyrProCysArgProAspValPheIle 587
 DB 3235 GGAAATGTACATTTTGGTGAAGTTGTATTCACTATCCCAACCCGACCGACATCCAGTG 3294
 QY 588 LeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSer 607
 DB 3295 CTTCAGGGACTGAGCTGGAGGTGAAGAGGGCCAGACCTGGCTCTGTGGGCAGCAGT 3354
 QY 608 GlyCysGlyLysSerThrSerValGlnLeuGlnArgLeuTyrAspProValGlnGly 627
 DB 3355 GGTGTGGGAAGACACAGTGTCTCGAGCTCTGGAGCGGTTCTACGACCCCTTGGCAGGG 3414
 QY 628 GlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGln 647


```

Db 3415 AAAGTCTGCTGATGGCAAGAAATAAAGCGACTGAATGTTTCAGTGGCTCCGAGCACAC 3474
QY 648 IleAlaIleValProGlnGluProValLeuPheAsnCySerIleAlaGluAenIleAla 667
Db 3475 CTGGGCACTGCTGCCAGGAGCCCATCTGTTGACTGCAGCATTCCTGAGAACATTCGCC 3534
QY 668 TyrGlyAspAsnSerArgValValProLeuAspGluIleIleIleIleIleIleIleIle 687
Db 3535 TATGGAGACAACAGCCGGGTGGTGCACAGGAAGAGATTTGTAGGGCAGCAAGGAGGCC 3594
QY 688 AsnIleHisSerPheIleGluGlyLeuProGluIleIleIleIleIleIleIleIleIle 707
Db 3595 ACATACATCGCTTCATCGAGTCACCTGCCTTAATAATATAGCACTAAAGTAGGACAAA 3654
QY 708 GlyAlaGlnLeuSerGlyGlyGlnIleIleIleIleIleIleIleIleIleIleIleIle 727
Db 3655 GGAACCTCAGCTCTCTGTCGCCCAACAAACAGCATTCGCCATAGCTGCTGCTGTTTGA 3714
QY 728 LysProIleIleLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluIle 747
Db 3715 CAGCCTCATATTTGCTTTGGATGAAGCCACGTCAGCTCTGGATACAGAAAGTGAAG 3774
QY 748 ValValGlnHisAlaLeuAspIleAlaIleIleIleIleIleIleIleIleIleIleIle 767
Db 3775 GTTGTCACAGACCTTGACAAAGCCAGAGAGCGGCGACCTGCATTTGATGCTCAC 3834
QY 768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyIleIleIle 787
Db 3835 CGCCTGTCACCATCCAGAAATGCAGACTTAATAGTGTGTTCCAGAATGCAGAGTCAAG 3894
QY 788 GluGlnGlyThrHisGlnGluLeuArgAsnArgAspIleIleIleIleIleIleIleIle 807
Db 3895 GAGCATGGCAGCATCAGCAGCTGCTGGCACAGAAAGGATCTATTTTCAATGGTCACT 3954
QY 808 AlaGlnSer 810
Db 3955 GTCCAGGCT 3963

RESULT 2
US-08-784-649A-5
; Sequence 5, Application US/08784649A
; Patent No. 5830697
; GENERAL INFORMATION:
; APPLICANT: Sikic, Branimir I
; TITLE OF INVENTION: P-GLYCOPROTEIN MUTANT RESISTANT TO
; TITLE OF INVENTION: CYCLOSPORIN MODULATION
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 2200 Sand Hill Road
; CITY: Menlo Park
; STATE: CA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/784,649A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: Reg.No. 5830697 36,677
; REFERENCE/DOCKET NUMBER: 06037/007001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-322-5070
; TELEFAX: 415-854-0875
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:

```

```

; LENGTH: 4264 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-784-649A-5

Alignment Scores:
Pred. No.: 6,88e-275 Length: 4264
Score: 2436.00 Matches: 466
Percent Similarity: 76.79% Conservativeness: 166
Best Local Similarity: 56.62% Mismatches: 177
Query Match: 59.72% Indels: 14
DB: 2 Gaps: 4

US-09-873-409-2 (1-812) x US-08-784-649A-5 (1-4264)
QY 2 ValAspGluAsnAspIleArgAlaLeuAenValArgHisTyrArgAspHisIleGlyVal 21
Db 1495 GTTGATGGACAGATATTAGGACCATTAATGTAAGTTTCTACGGGAATCATTTGTGTG 1554
QY 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
Db 1555 GTGAGTCAGAACTGTATTGTTGCCACACGATAGCTGAAAACATTCGCTATGCGCGT 1614
QY 42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
Db 1615 GAAATGTCCACATGATGAGATTGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTT 1674
QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyGlyGlyGlyGlyGlyGly 81
Db 1675 ATCATGAACCTCCCTCATAAATTTGACACCCCTGGTGGAGAGAGGGGCCCGCATTT 1734
QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
Db 1735 GTTGGCAGAAAGCAGAGATCGCCATTTGACCTGCTGCTGCTGCTGCTGCTGCTGCT 1794
QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
Db 1795 CTGCTGGATGAGGCCACGTCAGCTTGGACACAGAAAGCAAGCAGTGGTTCAGGTGGCT 1854
QY 122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
Db 1855 CTGGATAAGGCCAGAAAGGTCCGACCATTTGTATAGCTCATCGTTTGTCTACAGT 1914
QY 142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
Db 1915 CGTAATGCTGACGTCACTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1974
QY 162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln----- 178
Db 1975 GATGAACCTCATGAAAGAGAAAGGCAATTTACTTCAAACTTGTCAATATGCAGACAGCA 2034
QY 179 -----AspIleLysAlaAspGluGluGlnMetGluSerMetThrTyrSerThrGlu 195
Db 2035 AATGAAGTTGAATAGAAAATGCAGCTGATGAATCAAAAGTGAAATGATGCTGCTGGA 2094
QY 196 ArgLysThrAsnSerLeuProLeuHisSerVal-----LysSerIle 209
Db 2095 ATGCTCTTCAATGATTCAAGATCCAGTCTAATAAGAAAAGATCAACTCGTAGGAGTGTC 2154
QY 210 Lys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeu 228
Db 2155 CTGGATACACAGCCCAACAGCAAGAGCTTAGTACCAAGAGGCTCTGGATGAAGATATA 2214
QY 229 ProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTyrProPheValVal 248
Db 2215 CCTCCAGTTTCTTTTGGAGGATATGAAGCTAAATTTAACTGAATGGCTTATTGTTGT 2274
QY 249 LeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePhe 268
Db 2275 GTTGGTGTATTTTGTGCCATTATAAATGAGGCGCTGCAACACAGCATTTTGCATAATAT 2334

```


Qy	269	Ala	ys	lle	lle	e	Thr	Met	Phe	Gly	Asn	---	Asn	Asp	Ly	s	Thr	Thr	Leu	Leu	His	Asp	Ala	287
Db	2335	TC	AA	GA	AT	T	T	AG	GG	T	T	T	T	T	CA	CA	GA	A	T	T	G	A	A	2394
Qy	288	Gl	u	le	Tyr	Ser	Met	I	le	Phe	Val	I	le	Leu	Gly	Val	I	le	Cys	Phe	Val	Ser	Tyr	307
Db	2395	AA	CT	G	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	2454	
Qy	308	Gln	Gly	Leu	Phe	Tyr	Gly	Arg	Ala	Gly	I	le	Leu	Thr	Met	Arg	Leu	Arg	His	Leu	Ala	327		
Db	2455	C	A	G	G	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	2514		
Qy	328	Phe	Ly	s	Ala	Met	Leu	Tyr	Gln	Asp	I	le	Ala	T	Trp	Phe	Asp	Glu	Ly	s	Glu	Asn	347	
Db	2515	T	T	C	C	G	A	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	2574	
Qy	348	Gly	Leu	Thr	Thr	I	le	Leu	Ala	I	le	Asp	I	le	Ala	Gln	I	le	Gln	Gly	Val	Ala	367	
Db	2575	G	C	A	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	2634	
Qy	368	I	le	Gly	Val	Leu	Thr	Gln	Asn	Ala	Thr	Asn	Met	Gly	Leu	Ser	Val	I	le	I	le	Ser	387	
Db	2635	C	T	T	G	C	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	2694	
Qy	388	Tyr	Gly	Trp	Glu	Met	Thr	Phe	Leu	I	le	Leu	Ser	I	le	Ala	Pro	Val	Leu	Val	Ala	Thr	407	
Db	2695	T	A	T	G	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	2754	
Qy	408	Met	I	le	Glu	Thr	Ala	Ala	Met	Thr	Gly	Phe	Ala	Asn	Ly	s	Asp	Ly	s	Gln	Gly	Leu	427	
Db	2755	G	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	2814	
Qy	428	Ala	Gly	Ly	s	I	le	Ala	Thr	Glu	Ala	Leu	Glu	Asn	I	le	Arg	Thr	I	le	Val	Ser	447	
Db	2815	G	C	T	G	G	A	G	A	T	G	C	T	A	G	A	A	A	A	A	A	A	A	2874
Qy	448	Glu	Ly	s	Ala	Phe	Glu	Gln	Met	Tyr	Glu	Glu	Met	Leu	Gln	Thr	Gln	His	Arg	Asn	Thr	Ser	467	
Db	2875	G	A	C	A	G	A	A	G	T	T	T	T	T	T	T	T	T	T	T	T	T	2934	
Qy	468	Ly	s	Ly	s	Ala	Gln	I	le	I	le	Gly	Ser	Cys	Tyr	Ala	Phe	Ser	His	Ala	Phe	I	487	
Db	2935	A	G	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	2994	
Qy	488	Tyr	Ala	Ala	Gly	Phe	Arg	Phe	Gly	Ala	Tyr	Leu	I	le	Gln	Ala	Gly	Arg	Met	Thr	Pro	Glu	507	
Db	2995	T	A	T	C	T	G	A	T	T	T	T	T	T	T	T	T	T	T	T	T	T	3054	
Qy	508	Gly	Met	Phe	I	le	Val	Phe	Thr	Ala	I	le	Ala	Tyr	Gly	Val	Ala	Met	Ala	I	le	Gly	527	
Db	3055	G	A	T	C	T	C	T	G	T	A	T	T	T	T	T	T	T	T	T	T	T	3114	
Qy	528	Val	Leu	Ala	Pro	Glu	Tyr	Ser	Ly	s	Ala	Ly	s	Ser	Gly	Ala	Ala	His	I	le	Phe	Ala	547	
Db	3115	T	C	A	T	T	T	C	T	C	T	A	T	G	A	T	T	T	T	T	T	T	3174	
Qy	548	Glu	Ly	s	Ly																			

```

Db 3415 AAAGTGTCTTGTATGGCAAGAAATAAAGCAGCTGAATGTTCTAGTGGCTCCGAGCACAC 3474
Qy 648 IleaIaIleValProGlnGluProValLeuPheAsnCysSerIleAaGluAsnIleAa 667
Db 3475 CTGGGCATCGTGTCCAGAGGCCATCTCTGTTGACTGCAGCAITTCGTGAGAACATTTGCC 3534
Qy 668 TytGlyAspAsnSerArgValValProLeuAspGluIleLysGluAaIaAsnAaIa 687
Db 3535 TATGGAGACAACAGCCGGTGTGTACAGGAAGAGATTGTGAGGGCAGCAAAAGGAGGCC 3594
Qy 688 AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLys 707
Db 3595 AACATACATCCCTTCATCAGTCACTGGCTTAATAAATATAGCACTAAAGTAGGAGACAA 3654
Qy 708 GlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAaArgAlaLeuGln 727
Db 3655 GGAACACTCAGCTCTCTGGTGGCCAGAAACAACGCAATTCAGCTCTGTCCTGTTTGA 3714
Qy 728 LysProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 747
Db 3715 CAGCCTCATATTTTGTCTTTTGGATGAAGCCAGCTCAGCTCTGGATACAGAAAGTGAAAG 3774
Qy 748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767
Db 3775 GTTGTCCAGAAGCCCTGGACAAGCCAGAGAAGCCGCACTGCATTGTGATTGCTCAC 3834
Qy 768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys 787
Db 3835 CGCCTGTCCACATCCAGATGCAGACTTAATAGTGGTGTTCAGATGGCAGAGTCAAG 3894
Qy 788 GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807
Db 3895 GAGCATGGCAGCATCAGCAGCTGTCTGGCACAGAAAGGCATCTATTTTCAATGGTCAGT 3954
Qy 808 AlaGlnSer 810
Db 3955 GTCCAGGCT 3963

RESULT 3
US-08-181-471-2
; Sequence 2, Application US/08181471
; Patent No. 5641508
; GENERAL INFORMATION:
; APPLICANT: Li, Lingna
; APPLICANT: Lishko, Valeryi K.
; TITLE OF INVENTION: METHOD FOR DELIVERING BENEFICIAL
; TITLE OF INVENTION: COMPOSITIONS TO HAIR FOLLICLES
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Thomas Fitting
; STREET: 12526 High Bluff Drive, Suite 300
; CITY: San Diego
; STATE: CA
; COUNTRY: USA
; ZIP: 92130
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/181,471
; FILING DATE: 13-JAN-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/041,553
; FILING DATE: 02-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: ANT0029P
; TELECOMMUNICATION INFORMATION:

```


; TELEPHONE: 619-792-3680

; TELEFAX: 619-792-8477

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 4646 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; FEATURE: NO

; NAME/KEY: CDS

; LOCATION: 425..4267

; US-08-181-471-2

Alignment Scores:

Pred. No.: 8e-275 Length: 4646

Score: 2436.00 Matches: 466

Percent Similarity: 76.79% Conservative: 166

Best Local Similarity: 56.62% Mismatches: 177

Query Match: 59.72% Indels: 14

DB: 1 Gaps: 4

US-09-873-409-2 (1-812) x US-08-181-471-2 (1-4646)

QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21

DB 1781 GTTGATGGACAGATATTAGGACCATTAATGTAAAGTTTCTACGGGAATCATTTGGTGTG 1840

QY 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleIysTyrGlyArg 41

DB 1841 GTGAGTCAGGAACCTGTAATTTGTTCACACCATAGCTGAAACATTCGCTATGGCCGT 1900

QY 42 AspAspValThrAspGluMetGluArgAlaAargGluAlaAsnAlaTyrAspPhe 61

DB 1901 GAAATGTCAACATGATGAGATGTAGAAGCTGTCAAGGAAGCCATGCTATGACTTT 1960

QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81

DB 1961 ATCATGAACCTGCTCATAAATTTGACCCCTGGTTGGAGAGAGGGCCCAAGTTGAGT 2020

QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProIysIleLeu 101

DB 2021 GGTGGGACAGACAGAGGATCGCAATTCGACGTGCTGGTTCCCAACCCCAAGATCCTC 2080

QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerIysSerAlaValGlnAlaAla 121

DB 2081 CTGCTGGATGAGGCCACGTCAGCTTTGGACACAGAAGCGAGCGTTCAGGTGGCT 2140

QY 122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141

DB 2141 CTGGATAAGCCAGAAAAGTCCGACCAACATTTGTATAGTCTCATCGTTTGCTACAGTT 2200

QY 142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161

DB 2201 CGTAATGCTGACGTCACTCGTGGTTTCGATGATGGAGTCATTCGGAGAAAGGAATCAT 2260

QY 162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerIleuValMetSerGln----- 178

DB 2261 GATGAATCATGAAGAGAAAAGGCATTTACTTCAAACTTTGTCAAAATGACAGACAGGGA 2320

QY 179 -----AspIleLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu 195

DB 2321 AATGAAGTTGAATTAGAAATGACGTGATGAATCCAAAGTCAAAATGATGATGCTTGGAA 2380

QY 196 ArgLysThrAsnSerLeuProLeuHisSerVal-----LysSerIle 209

DB 2381 ATGCTCTCAATGATTTCAAGATCCAGTCTTAATAGAAAAGATCAACTCGTAGGAGTGTC 2440

QY 210 Lys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeu 228

DB 2441 CGTGGATCAACGCCCAAGACAGAAAGCTTAGTACCAGAGAGGCTCTGGATGAAGATATA 2500

QY 229 ProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTyrProPheValVal 248

DB 2501 CCTCCAGTTTCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGGCTTATTGTGT 2560

QY 249 LeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePhe 268

DB 2561 GTTGGTGTAATTTGTGCCATTATAAATGGAGGCTGCACACAGCATTTGCAATAATATT 2620

QY 269 AlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeuLysHisAspAla 287

DB 2621 TCNAAATTTATAGGGGTTTTTACAGAATTTGATGATCTCGAAACAAACACAGATAGT 2680

QY 288 GluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMet 307

DB 2681 AACTTGTCTTCACTATTGTTCTAGCCCTTGAATATTATTTCTTTTATACATTTTCCCT 2740

QY 308 GlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla 327

DB 2741 CAGGGTTTCACATTTGGCAAGCTGGAGAGATCCTCACCAAGCGCTCCGATACATGGT 2800

QY 328 PheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGluLysGluAsnSerThrGly 347

DB 2801 TTCCGATCCATGCTCAGACAGATGTGAGTTGGTTGATGACCTTAAACACCATGGA 2860

QY 348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg 367

DB 2861 GCATTGACTACCAAGCTCGCAATGATGCTGCTCAAGTTAAAGGGCTATAGTTCAGG 2920

QY 368 IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387

DB 2921 CTTGCTGTATATACCAGAAATATAGCAAAATCTTGGACAGGAATAATATATATCTTCATC 2980

QY 388 TyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaproValLeuAlaValThrGly 407

DB 2981 TATGGTTGCAACTTAACACTGTCTTAGCAATTTGTACCCATCATTTGCAATAGCAGGA 3040

QY 408 MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHis 427

DB 3041 GTTGTGAAATGAAATGTTGCTGCAACAGCACTGAAAGATAAGAAAGAACTAGAAAGT 3100

QY 428 AlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArg 447

DB 3101 GCTGGGAAGATCGCTACTGAAGCAATAGAAAACCTTCGAAACCGTTCTTTCTTTGACTCC 3160

QY 448 GluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSer 467

DB 3161 GACGAGAAGTTTGAACATATGTATGCTCAGAGTTTCAGGTACCACACAGAAACTCTTTG 3220

QY 468 LysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487

DB 3221 AGGAAGACACATCTTTTGAATTTACATTTTCTTCACCCAGGCAATGATGATTTTTC 3280

QY 488 TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507

DB 3281 TATGCTGGATGTTTCGGTTTGGAGCTACTTGGTGGCACATAAACTCATGAGCTTTGAG 3340

QY 508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyValaMetAlaIleGlyLysThrLeu 527

DB 3341 GATGTTCTGTAGTATTTTTCAGCTGTTGCTTTGGTGCCATGGCCGTGGGGCAAGTCAGT 3400

QY 528 ValLeuAlaProGluTyrSerLysAlaLysSerGlyValaAlaHisLeuPheAlaLeuLeu 547

DB 3401 TCATTTGCTCTGACTATTCGCAAGCCAAATAATATCAGACAGCCACATCATCATGATCAT 3460

QY 548 GluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGlu 567

DB 3461 GAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGGCTTAATGCCGAACACATTTGGA 3520

QY 568 GlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIle 587

DB 3521 GGAAATGTCACATTTTGGTGAAGTTGTAATTCACATATCCCAACCCGACCGGACATCCCGAG 3580

Qy 328 PheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGly 347
Db 2801 TTCCGATCCATGCTCAGACAGGATGTGGTGTGATGAGCCCTCAAAAACACCATGGA 2860
Qy 348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg 367
Db 2861 GCATTGATACAGGCTCCCAATGATGCTGCTCAAGTTAAAGGGCTATAGGTTCAGG 2920
Qy 368 IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387
Db 2921 CTTCGTGTAATCCAGAAATATAGCAATCTTGGACAGGAATAATATATATCTTCATC 2980
Qy 388 TyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGly 407
Db 2981 TATGTTGGCAACTAACACTGTTACTCTTAGCAATTTGACCCATCTTGTCAATGACAGG 3040
Qy 408 MetIleGluThrAlaIleMetThrClyPheAlaAsnLysAspLysGlnIleuLysHis 427
Db 3041 GTTGTGTAATGAAATGTTGTGGCAAGCAGCTGAAAGATAGAAAGAACTAGAAGGT 3100
Qy 428 AlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArg 447
Db 3101 GCTGGAGAGATGCTACTGAGCAATAGAAACTTCCGACCGTGTGTTCTTGACTCAG 3160
Qy 448 GluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSer 467
Db 3161 GAGCAGAAAGTTGAACATATGATGCTCAGAGTTTGCAGGTACCATACAGAAACTCTTTG 3220
Qy 468 LysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487
Db 3221 AGGAAAGCACACATCTTTGGAAATATACATTTCTTCCTCCACCGGCAATGATGATATTTTCC 3280
Qy 488 TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507
Db 3281 TATGCTGATGTTCCGGTTTGGAGCCTACTTGTGGGCACATAAACTCATGAGCTTTCAG 3340
Qy 508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeu 527
Db 3341 GATGTTCTGTTAGTATTTTCAGCTTCTGCTTGTGTCATGGCCGTCGGGCAAGTCAGT 3400
Qy 528 ValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeu 547
Db 3401 TCATTGCTCTGCTGATATGCCAAAGCCAAAATATATCAGCAGGCCACATCATCATGATCAT 3460
Qy 548 GluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGlu 567
Db 3461 GAAAGAAACCCCTTGTATGACAGCTACAGCAGGAGGCTTAATCCCAACACATTGGA 3520
Qy 568 GlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIle 587
Db 3521 GGAATGTGCACATTTGGTGAAGTTGTATTCAACTATCCACCACCGACGACATCCAGTG 3580
Qy 588 LeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSer 607
Db 3581 CTTCAGGAGCTGAGCTGAGGTGGAAGGCGCCAGACGCTGCTGCTGGGCGAGCAGT 3640
Qy 608 GlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGly 627
Db 3641 GCCTGTGGAGAGACAGACAGTGTGTCAGCTCTCTGGAGCGGTTCTACAGCCCTTGCAGGG 3700
Qy 628 GlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGln 647
Db 3701 AAAGTGTCTTGTATGGCAAGAAATAAAGCGACTGAATGTTTCAGTGGCTCCGAGCACAC 3760
Qy 648 IleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAla 667
Db 3761 CTGGGCATGCTGTCCAGGAGCCCATCTGTTGATCTGAGCGGTTCTACAGCCCTTGCAGGG 3820
Qy 668 TyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAla 687
Db 3821 TATGAGACAAACGCGGGTGTGTGCACAGAAAGATGCTGAGGCGGCAAGAGGAGCC 3880
Qy 688 AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLys 707

Db 3881 AACATACATGCTTCATCGAGTCACTGCCCTAAATAATATAGCTAAAGTAGGAGACAAA 3940
Qy 708 GlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGln 727
Db 3941 GGAACCTCAGCTCTCTGCTGGCCAGAAACACGCAATGCCATAGCTCGTGCCTTGTAGA 4000
Qy 728 LysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 747
Db 4001 CAGCCTCATATTTTGTCTTTGGATGAAGCCACGTCAGCTCTGGATACAGAAAGTGAAG 4060
Qy 748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767
Db 4061 GTTGTCCAAAGAGCCCTGGACAAAGCCAGAGAGCCGACCTGCATTTGATTGCTCTC 4120
Qy 768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys 787
Db 4121 CGCTGTCCACCATCCAGATGCAGACTTAATAGTGTGTTTCAGATGGCAGAGTCAAG 4180
Qy 788 GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807
Db 4181 GAGCATGGCAGCATCAGCAGCTCTGGCACAGAAAGGCATCTATTTTCAATGGTCAGT 4240
Qy 808 AlaGlnSer 810
Db 4241 GTCCAGGCT 4249
RESULT 5
US-08-793-610-5
; Sequence 5, Application US/08793610
; Patent No. 5858744
; GENERAL INFORMATION:
; APPLICANT: BAUM, Christopher
; APPLICANT: STOCKING-HARBERS, Carol
; APPLICANT: OSTERTAG, Wolfram
; TITLE OF INVENTION: RETROVIRAL VECTOR HYBRIDS AND THE USE THEREOF
; TITLE OF INVENTION: FOR GENE TRANSFER
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nikolaide, Marmelstein, Murray & Oram LLP
; STREET: 655 Fifteenth Street N.W. Suite 330
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/793,610
; FILING DATE: 07-MAR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 44 31 973.8
; FILING DATE: 08-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 195 03 952.1
; FILING DATE: 07-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP95/03175
; FILING DATE: 10-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Berman, Richard J.
; REGISTRATION NUMBER: 39,105
; REFERENCE/DOCKET NUMBER: F1614-7007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)638-5000
; TELEFAX: (202)638-4810
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6505 base pairs
; TYPE: nucleic acid

STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA
US-08-793-610-5

Alignment Scores:

Pred. No.: 1.45e-274 Length: 6505
Score: 2436.00 Matches: 466
Percent Similarity: 76.79% Conservative: 166
Best Local Similarity: 56.62% Mismatches: 177
Query Match: 59.72% Indels: 14
DB: 2 Gaps: 4

US-09-873-409-2 (1-812) x US-08-793-610-5 (1-6505)

QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
DB 3173 GTTGATGGACAGATATTAGACCAATAATGTAAGTTTCTACGGGAATCATTTGGTGG 3232
QY 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
DB 3233 GTGAGTCAGAACTGTATTGTTTGGCCACCATAGCTGMAAACATTCGCTATGGCCGT 3292
QY 42 AspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
DB 3293 GAAATGTCCACCATGATGAGATTGAGAAAGCTGTCAAGGAAGCCCAATGCCATGACTTT 3352
QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
DB 3353 ATCATGAACTGCTCATTAATTTTGACCCCTGGTTGGAGAGAGGGGCCCATTTGAGT 3412
QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
DB 3413 GGTGGGCAGAAAGCAGAGGATCGCATTTGCACGTGCTGTTTCGCAACCCCAAGATCCTC 3472
QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
DB 3473 CTGCTGGATGAGGCCACGTCAGCTTTGGACACAGAAAGCGAGCAGTGGTTCAGGTGGCT 3532
QY 122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
DB 3533 CTGGATAGCCCAAGAAAGTCGGACACCATTTGTATAGCTCATCTGTTTGTCTACAGTT 3592
QY 142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
DB 3593 CGTAATGCTGACGTCATCGCTGTTTCGATGATGGAGTCATTGTGGAGAAAGGAATCAT 3652
QY 162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln----- 178
DB 3653 GATGAATCATGAAGAGAAAGACATTTACTTCAAACTGTTCACAATGCAGACAGCAGGA 3712
QY 179 -----AspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu 195
DB 3713 AATGAAGTTGAATTAGAAATGCGATGATGAATCCAAAGTCCAAATGATGATGCTTGGAA 3772
QY 196 ArgLysThrAsnSerLeuProLeuHisSerVal-----LysSerIle 209
DB 3773 ATGCTCTCAATGATTCAGATCCAGTCTAATAAGAAAGAAAGATCAACTCGTAGGAGTGC 3832
QY 210 Lys---SerAspPheIleAspLysAlaGluSerThrGlnSerLysGluIleSerLeu 228
DB 3833 CGTGGATCAACGCCCAAGACAGAAAGCTTAGTACCAGAGAGCTCTGGATGAAGATATA 3892
QY 229 ProGluValSerLeuLysIleLeuLysLeuAsnLysProGluIleThrProPheValVal 248
DB 3893 CTTCCAGTTCTTTTGGAGGATATGAAGCTAAATTTAACTGAATGGCCCTTATTGTT 3952
QY 249 LeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePhe 268
DB 3953 GTTGGTGTATTTGTGCCATTATAAATGGAGGCTGCAACCCAGCATTTGCAATAATATT 4012
QY 269 AlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeuLysHisAspAla 287
DB 4012

DB 4013 TCAAGATTATAGGGGTTTTTACAAGAATTGATGATCTCTGAACAAAACACAGCAATAGT 4072
QY 288 GluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMet 307
DB 4073 AACTTGTCTTCTACTATTGTTCTAGCCCTTGGAATATTCTTTTATTACATTTTCTCT 4132
QY 308 GlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla 327
DB 4133 CAGGGTTTTCACATTTGGCAAGCTGGAGAGATCCTCACCAGCGGCTCGGATACATGTT 4192
QY 328 PheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGluLysGluAsnSerThrGly 347
DB 4193 TTCCGATCCATGCTCAGACAGGATGTGAGTTGTTGTATGACCCCTAAACACCACTGGA 4252
QY 348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg 367
DB 4253 GCATTGACTACCGCTCGCCCAATGATGCTGCTCAAGTTAAAGGGGCTATAGTTCACAG 4312
QY 368 IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387
DB 4313 CTTGCTGTAATTAACCCAGAAATATAGCAAACTCTGGGACAGGAATAATATATCTCTCATC 4372
QY 388 TyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGly 407
DB 4373 TATGGTTGGCAACTAACACTGTTACTCTTAGCAATTTGACCATCATTCGAATAGCAGGA 4432
QY 408 MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLysHis 427
DB 4433 GTTGTGAAATGAAATGTTGCTGGCAAGCACTGAAAGATAAGAAAGAACTAGAGGT 4492
QY 428 AlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArg 447
DB 4493 GCTGGGAAGATCGCTACTGAAGCAATAGAAAATCTCCGAACCGTTGTTCTTTGACTCAG 4552
QY 448 GluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSer 467
DB 4553 GAGCAGAAGTTTGAACATATGATGCTCAGAGTTTGCAGGTACCATACAGAAACCTTTTG 4612
QY 468 LysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487
DB 4613 AGGAAGACACACATCTTTGGAATTACATTTCTTCCACCGCAATGATGATTTTCTCC 4672
QY 488 TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507
DB 4673 TATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAAACTCATGAGCTTGG 4732
QY 508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeu 527
DB 4733 GATGTTCTGTAGTATTTTCAGCTGTGTTGTTGTCCTGCGCCGTGGGCAAGTCAGT 4792
QY 528 ValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeu 547
DB 4793 TCATTTGCTCTGACTATGCCAAAGCCAAATAATATCAGACGCCCATCATCATGATCAT 4852
QY 548 GluLysLysProAsnIleAspSerArgSerGlnGluGlyLysPheAspThrCysGlu 567
DB 4853 GAAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGGCTTAATGCCGAACACATTTGGA 4912
QY 568 GlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIle 587
DB 4913 GGAATGTCATCATTTGGTGAATTTGTATTCAACTATCCACCCGACCGACATCCAGTG 4972
QY 588 LeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSer 607
DB 4973 CTTCAGGACTGAGCTGGAGGTGAAGAGGCCAGACGCTGCTCTGTTGGGCGACAGT 5032
QY 608 GlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGly 627
DB 5033 GGCTGTGGGAAGACAGCACAGTGTCTCAGCTCTCTGGAGCGGTTCTACGACCCCTTTGG 5092
QY 628 GlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGln 647
DB 5093 AAGATGCTCTCTTGTGGGCAAGAAATAAAGCCGACTGAATGTTTCAGTGGTCCGACGAC 5152


```
Qy 648 IleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAla 667
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 5153 CTGGGCATCGTCTCCAGGAGCCCATCTCTGTTGACTGCAGCATTTGCTGAGAACATTGCC 5212
Qy 668 TyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAla 687
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 5213 TATGAGACAACAGCCGGTGGTGTCAACAGGAAGAGATCGTGAGGCGCAGCAAGAGGCC 5272
Qy 688 AsnIleHisSerPheIleGluGluLeuProGluLysTyrAsnThrGlnValGlyLeuLys 707
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 5273 AACATACATGCTTCATCGAGTCATCGCTCTAATAATATAGCACTAAAGTAGGACACAA 5332
Qy 708 GlyAlaGlnLeuSerGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGln 727
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 5333 GGAACCTCAGCTCTCTGGTGGCCAGAAACAAACGCAATGGCCATAGCTCGTGCCCTTGTAGA 5392
Qy 728 LysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnSerGluLys 747
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 5393 CAGCCTCATATTTTCTTTGGATGAAGCCACGTCAGCTCTGGATACAGAAAGTGAAGAAG 5452
Qy 748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyAlaGThrCysLeuValValThrHis 767
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 5453 GTTGTCGAAGAGCCCTGGCAAAACCCAGAGAGAGCCGACCTGCTGTTGTGATTCCTCAC 5512
Qy 768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys 787
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 5513 CGCCTGTCACCATCCAGAAATCCAGACTTAATAGTGGTGTTCAGAAATGGCAGAGTCAAG 5572
Qy 788 GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 5573 GAGCATGGCAGCCATCAGCAGCTGCTGGCAGAGAGGATCTATTTTCAATGCTCAGT 5632
Qy 808 AlaGlnSer 810
   : : : : :
Db 5633 GTCCAGGCT 5641

RESULT 6
US-08-793-610-6
; Sequence 6, Application US/08793610
; Patent No. 5858744
; GENERAL INFORMATION:
; APPLICANT: BAUM, Christopher
; APPLICANT: STOCKING-HARBERS, Carol
; APPLICANT: OSTERTAG, Wolfram
; TITLE OF INVENTION: RETROVIRAL VECTOR HYBRIDS AND THE USE THEREOF
; TITLE OF INVENTION: FOR GENE TRANSFER
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nikaido, Marmelstein, Murray & Oram LLP
; STREET: 655 Fifteenth Street N.W. Suite 330
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/793,610
; FILING DATE: 07-MAR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 44 31 973.8
; FILING DATE: 08-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 195 03 952.1
; FILING DATE: 07-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP95/03175
; FILING DATE: 10-AUG-1995
; ATTORNEY/AGENT INFORMATION:
```

```
; NAME: Berman, Richard J.
; REGISTRATION NUMBER: 39,105
; REFERENCE/DOCKET NUMBER: P1614-7007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)638-5000
; TELEFAX: (202)638-4810
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9318 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA
; US-08-793-610-6

Alignment Scores:
Pred. No.: 2,72e-274 Length: 9318
Score: 2436.00 Matches: 466
Percent Similarity: 76.79% Conservative: 166
Best Local Similarity: 56.62% Mismatches: 177
Query Match: 59.72% Indels: 14
DB: 2 Gaps: 4
```

US-09-873-409-2 (1-812) x US-08-793-610-6 (1-9318)

```
Qy 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 3132 GTTGATGACAGGATATTAGGACCATTAATGTAAGTTTCTACGGGAAATCATTTGCTGTG 3191
Qy 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 3192 GTGAGTCAGGAACCTGTATTGTTTCCACACCATGATGAAACATTCGCTATGCCGCT 3251
Qy 42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 3252 GAAATGTCACCATGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAATGCCATGACITT 3311
Qy 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyValAlaGlnMetSer 81
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 3312 ATCATGAAATGCCTCATAAATTTGACACCCCTGGTTGGAGAGAGAGGGGCCAGTTGAGT 3371
Qy 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 3372 GGTGGGAGAGACAGAGGATGCCATTCACGTGCCCTCGCTGGTTCGCAACCCCAAGATCCTC 3431
Qy 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 3432 CTGCTGGATGAGCCACGTCAGCCTTGACACAGAAACGAGCAGTGGTTCAGGTGGCT 3491
Qy 122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaAlaHisArgLeuSerThrIle 141
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 3492 CTGGATAAGGCCAGAAAGGTCGGACCCACCATTTGTGATAGCTCATCGTTGTCTACAGTT 3551
Qy 142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 3552 CGTAATGCTGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGCGAAGAAAGAAATCAT 3611
Qy 162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln----- 178
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 3612 GATGAACATCATGAAGAGAAAGAGGCAATTTACTTCAAACTGTGCACAAATGCAGACAGGA 3671
Qy 179 -----AspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu 195
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 3672 AATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAATGATGATCCCTTGAA 3731
Qy 196 ArgLysThrAsnSerLeuProLeuHisSerVal-----LysSerIle 209
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 3732 ATGCTCTCAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACTCGTAGGAGTGTG 3791
Qy 210 Lys----SerAspPheIleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeu 228
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 3792 CGTGGATCAACAGCCCAAGAGAGAGGCTTAGTACCAAGAGGCTCTCGATGAAAGTATA 3851
```


QY 229 ProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpPropheValVal 248
Db 3852 CTTCCAGATTTCTTTGGAGGATATGAGCTAAATTTAACTGAGCGCTTATTTGTT 3911
QY 249 LeuGlyThrLeuAlaSerValLeuAenGlyThrValHisProValPheSerIlePhe 268
Db 3912 GTTGGTGATTTTGTGCTATTAATATGAGGCTGCAACACGATTTGCAATATATTT 3971
QY 269 AlaLysIleLeuThrMetPheGlyAsn---AsnAspLysThrThrLeuLysHisAspAla 287
Db 3972 TCAAGATTTATAGGGTTTTTACAGAAATGATGATCCTGAAACAAACACAGAGATAGT 4031
QY 288 GluIleTyrSerMetLeuPheValIleLeuGlyValIleCysPheValSerTyrPheMet 307
Db 4032 AACTTGTGTTTCACTATTTGTTTCTAGCCCTTGAATTTATTTTATCATTTTCTCT 4091
QY 308 GlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla 327
Db 4092 CAGGGTTTCATTTGGCAAGCTGGAGAGATCCTCACCAGCGGCTCCGATACATGGTT 4151
QY 328 PheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGly 347
Db 4152 TTCCGATCCATGCTCAGACAGGATGTCAGTTGTTGATGACCCCTAAACACACCATGGA 4211
QY 348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg 367
Db 4212 GCATTCAGTACCAGGCTCGCAATGATGCTGCTCAAGTTAAAGGGCTATAGGTTCCAGG 4271
QY 368 IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387
Db 4272 CTTGCTGTAAATACCCAGATATAGCAAAATCTTGGACAGGAATAATATATCCTTCATC 4331
QY 388 TyrGlyTrpGluMetThrPheIleLeuSerIleAlaProValLeuAlaValThrGly 407
Db 4332 TATGGTTGGCACTAACACTGTTACTTCTAGCAATGTACCCATCATTCGATAGCAGGA 4391
QY 408 MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHis 427
Db 4392 GTTGTGTAATGAAATGTTGCTGGACAGCACTGAAAGATAAGAAAGAACTAGAAGGT 4451
QY 428 AlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArg 447
Db 4452 GCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCGAAACCGTTGTTCTTTGATCTCAG 4511
QY 448 GluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSer 467
Db 4512 GAGCAGAGTTTGAACATATGATGCTCAGAGTTTGCAGGTACCATACAGAAACTCTTTG 4571
QY 468 LysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487
Db 4572 AGGAAAGCACACATCTTTGGAATTTACATTTTCTTCCACCCAGGCAATGATGATTTTCC 4631
QY 488 TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507
Db 4632 TATGCTGGATGTTTCGGTTTGGAGCTACTTGTGGCACAATAAATCTCATGAGCTTTGAG 4691
QY 508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeu 527
Db 4692 GATGTTCTGTAGTATTTTCACTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 4751
QY 528 ValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeu 547
Db 4752 TCATTTGCTCTGCTATGCTATGCAAGCAAAATATATCAGCAGCCCATCATGATCAT 4811
QY 548 GluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGlu 567
Db 4812 GAAAAAACCCCTTTGATTTGACACTACAGACGGAAGGCTTAATGCGGACACATTTGGA 4871
QY 568 GlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIle 587
Db 4872 GGAATGTCACATTTGGTGAAGTTGTTATTCATCTCCACCCGACCGGACATCCCATG 4931
QY 588 LeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSer 607

Db 4932 CTTCCAGGACTCAGCTCGAGGTGAAGAGGCGCACGCTGCTCTGTGGCCAGCAGT 4991
QY 608 GlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGly 627
Db 4992 GCTGTGGGAAGAGCACAGTGTCTCAGCTCTCTGGAGCGTTCTACGACCCCTTGGCAGG 5051
QY 628 GlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGln 647
Db 5052 AAGTGTCTCTTGTATGGCAAGAAATAAAGCAGCTCAATGTTCACTGGCTCCGAGCAC 5111
QY 648 IleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAla 667
Db 5112 CTGGGATCGTCTCCAGGAGCCCATCTGTTGATCTGAGCATTTGCTGAGAACATTTGCC 5171
QY 668 TyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAla 687
Db 5172 TATGAGACAAACGCGGCTGTGTACAGGAAGAGATCTGTCAGGCGCAGCAAGGAGGCC 5231
QY 688 AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLys 707
Db 5232 AACATACATGCTTCTCATCGAGTCACTGCCTTAATAATATAGCACTAAAGTAGGACAAA 5291
QY 708 GlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuGln 727
Db 5292 GGAACCTCAGCTCTGCTGGTGGCCAGAAACACGATTTGCCATAGCTCGTCCCTTGTAGA 5351
QY 728 LysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 747
Db 5352 CAGCTCATATTTTCTTTGATGAGCCAGCTCAGCTCTGATACAGAAAGTGAAG 5411
QY 748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767
Db 5412 GTTGTCCAAAGAGCCCTGGACAAAGCCAGAGAGGCGCACCTGCACTTGATTTGCTCAC 5471
QY 768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValLeuHisGlnGlyLysIleLys 787
Db 5472 CGCTGTCTCCACCATCCAGAAATGCACTTAATAGTGTGTTTTCAGAAATGCGCAGAGTCAAG 5531
QY 788 GluGlnGlyThrHisGlnGluLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807
Db 5532 GAGCATGGACCATCAGCAGCTGTGGCAGCAAGGCAATCTATTTTCAATGTCAGT 5591
QY 808 AlaGlnSer 810
Db 5592 GTCCAGGCT 5600
RESULT 7
US-09-120-513-1
; Sequence 1, Application US/09120513
; Patent No. 6025160
; GENERAL INFORMATION:
; APPLICANT: Brun, Kimberly
; APPLICANT: Chenery, Richard
; APPLICANT: Ellens, Harma
; APPLICANT: Field, John
; APPLICANT: Yue, Lin
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE
; TITLE OF INVENTION: SEQUENCES ENCODING RAT MDR1B2 AND
; NUMBER OF INVENTION: SCREENING METHODS THEREOF
; CORRESPONDENCE ADDRESS:
; ADDRESS: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY:
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0


```

: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/120,513
: FILING DATE: 22-JUL-1998
:
: CLASSIFICATION:
:
: ATTORNEY/AGENT INFORMATION:
: NAME: King, William T
: REGISTRATION NUMBER: 30,954
: REFERENCE/DOCKET NUMBER: GP50008
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 610-270-5015
: TELEFAX: 610-270-5090
:
: TELEX:
:
: INFORMATION FOR SEQ ID NO: 1:
:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 4233 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
:
: MOLECULE TYPE: Genomic DNA
: US-09-120-513-1

```

Alignment Scores:			
Pred. No.:	2,976-272	Length:	4233
Score:	2413.50	Matches:	467
Percent Similarity:	76.01%	Conservative:	158
Best Local Similarity:	76.01%	Mismatches:	180
Query Match:	59.11%	Indels:	17
DB:	3	Gaps:	5
US-09-873-409-2 (1-812) x US-09-120-513-1 (1-4233)			
Qy	2	ValAspClnuAsnAspIleArgAlaLeuAenValArgHisTyrArgAspHisIleGlyVal	21
Db	1376	ATCGAGCGACAGACATCAGGACCATCATGTGAGGTATCGCGGGAATCATTTGGGGTG	1435
Qy	22	ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg	41
Db	1436	GTGAGTCAGGACCGTGCTGTTGGCCACACGATTCGCGAAACATTCGCTATGGCGGA	1495
Qy	42	AspAspValThrAspGluMetGluArgAlaIaArgGluAlaAsnAlaTyrAspPhe	61
Db	1496	GAAACGTCACCATGGATGAGATAGAGAAGCTGTCAGGAAGCCAAATGCCTATGACTTC	1555
Qy	62	IleMetGluPheProAnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer	81
Db	1556	ATCATGAAGTGCCTCCACAAATTAAACACCTCGTGTGTGAGAGGGGCGCAGCTAGT	1615
Qy	82	GlyGlyGlnLysGlnArgIleAlaIleAArgAlaLeuValAArgAsnProLysIleIeu	101
Db	1616	GGGGGACAGAAAACAGAGGATCGCATTCGCCGGGGCCCTGGTCCGCAACCCCAAGATCCT	1675
Qy	102	IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla	121
Db	1676	TTGTTGGATGAGCCACGTCAGCCTTGGACACAGAAAGCGGCGGTTCAGGCGCGCT	1735
Qy	122	LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle	141
Db	1736	CTGGATAAGGCTAGAGAAGCCGGACCAACCATTTGTATAGTCTACCGCTTGTCTACAGC	1795
Qy	142	ArgSerAlaAspIleuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis	161
Db	1796	CGCAATGCTCAGCTCATTCGCTGTTTTGTATGGTGTGTCATTGTTCGACCAAGGAATCAT	1855
Qy	162	AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIleLys	181
Db	1856	GAAGAGCTCATGAAGAGGAGGCAATTTACTTCAAACTTGTCTATGACACAG---ACTAGA	1912
Qy	182	LysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeu	201
Db	1913	GGAAATGAAATTGACCCAGGAATAATGCTTATGAATCCCAAGTGCACACTGGTGCCTCT	1972
Qy	202	ProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAla	217

Dbb 2210 TTTTCAAGATTGTAGGGGTTTTTTCAGAGACAGCAGGACCATGAACCAAAACCAACCGAAT 2269
Qy 287 AlaGluileTyrSerMetilePheValilleLeuGlyValilleCysPheValserTyrPhe 306
Db 2270 TGTAACTTGTTCCTTCCTCTCTTCTGGTTCATGGGAATGATTTCTTTTGTACGTACTTC 2329
Qy 307 MetGlnGlyLeuPheTyrGlyArgAlaGlyGluileLeuThrMetArgLeuArgHisLeu 326
Db 2330 TTTCAAGGCTTCACATTTTGGCAAGCTGGAGAGATCCTCACCAACGCACTCCGATACATG 2389
Qy 327 AlaPheLysAlaMetLeuTyrGlnAspilleAlaTyrPheAspGluLysGluAsnSerThr 346
Db 2390 GTCTTCAAAATCCATGCTGGCAGAGATATAGCTGTTTGTATGACCAATAAAACCACT 2449
Qy 347 GlyGlyLeuThrThreLeuAlaileAspilleAlaGlnileGlnileValathrGlySer 366
Db 2450 GGCTCGCTGACTACCAAGGCTCGTAGTACGCTTCTAATGTTTAAAGGGGCTATGGGCTCC 2509
Qy 367 ArgilleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValilleSerPhe 386
Db 2510 AGGCTTGTCTGTATACCCAGAAATGTAGCAAACTTGGCACAGGAATATATCTTATCCTTA 2569
Qy 387 IleTyrGlyTyrGluMetPheLeuileLeuSerIleAlaProValleuAlaValThr 406
Db 2570 GTCTATGGCTGGCAGCTTACACTTTTACTTGTAGTAATATATACCACTCAATTCCTTGGGT 2629
Qy 407 GlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLys 426
Db 2630 GGAATATTTGAATGAACTGTGTCTGCTCAAGCTTGAAGGACAAAGAAAGCTAGAG 2689
Qy 427 HisAlaGlyLysIleAlaThrGluAlaLeuGluAsnileAArgThrIleValserLeuThr 446
Db 2690 ATCTCTGGGAAGATCGCTACAGAAGCAATTTGAAACTTCCGCACTGTTGTCTCTTTGACT 2749
Qy 447 ArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThr 466
Db 2750 AGGGAGCAGAAGTTTTGAACATGATGATGCCAGAGCTTGCAGATACCACAGCAATGCT 2809
Qy 467 SerLysLysAlaGlnIleileGlySerCysTyrAlaPheSerHisAlaPheileTyrPhe 486
Db 2810 TTGAAGAAGCACACGCTCTTTGGGATCACCTTCGCTTCACCCAGGCCATGATTTATTTT 2869
Qy 487 AlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuileGlnAlaGlyArgMetThrPro 506
Db 2870 TCCTATGCTGCTTGTTCGGTTCGGTGTCTACTTGGTGGCCAGAACTCATGACGTTT 2929
Qy 507 GluGlyMetPheileValPheThrAlaileAlaTyrGlyAlaMetAlaileGlyLysThr 526
Db 2930 GAAATGTTATGTATGTTATTTCTGCTGTGTTTGTGGTGCATGGCAGCAGGAATACC 2989
Qy 527 LeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeu 546
Db 2990 AGTTTCATTCCTGCTGACTACGGAAGGCCAAAGTCTCAGCATCCCATCATCAGGATC 3049
Qy 547 LeuGluLysLysProAsnileAspSerArgSerGlnGluGlyLysLysProAspThrCys 566
Db 3050 ATTGAGAAAATCCCCAGATTTCAGACCTACAGCAGGAGGCTTGAAGCCTAATTGGTTA 3109
Qy 567 GluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPhe 586
Db 3110 GAAGGAAATGTGAAATTTAATGGAGTCATGTTCAACTATCCCAACCCAGCCCAACATCCA 3169
Qy 587 IleLeuArgGlyLysSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySer 606
Db 3170 GTGCTTCAGGACTGAGCTTCGAGGTGAGAGGGGCAACCGTTCGCTGTGGGCAGC 3229
Qy 607 SerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGln 626
Db 3230 AGTGGCTCGGGAAGAGTACAGTGGTCCAGCTCTCGAGCGCTTCTACAACCCCATGGCT 3289
Qy 627 GlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnThrLeuArgSer 646
Db 3290 GGAACAGTGTCTTAGTAGGCAAGAAATAAAACAACTCAACGCTCAGTGGCTCCGCGCC 3349

Qy 647 GlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIle 666
Db 3350 CACCTGGGCATTTGTTCAGAGCCCATCTGTTGTGACTCAGCATCACCGAGAACATC 3409
Qy 667 AlaTyrGlyAspAsnSerArgValValProLeuAspGluileLysGluAlaAlaAsnAla 686
Db 3410 GCCTACGAGAACACAGCCGTGTCTGTCTCATGAGGAGATCGTCAGGCGCCCGCAGGAG 3469
Qy 687 AlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeu 706
Db 3470 GCCAACATCCACAGTTTCATCGACTCACTGCTTGAGAAATACAAACACAGAGTGGGAGAC 3529
Qy 707 LysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaileAlaArgAlaLeuLeu 726
Db 3530 AAAGGCACTAGCTGTTCGGGGGGCAGAAAGCAGCGCATCGCGCCCGCTCGTTC 3589
Qy 727 GlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnSerGlu 746
Db 3590 AGACAGCTTCACATCTTACTTCTGATGAAGCAGCATCAGCTCTGGATACGAGAGTGA 3649
Qy 747 LysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThr 766
Db 3650 AAGTCTGTCAGGAAGCGCTGGACAAACCCAGGAAGCCGACCTGCTGATCGCG 3709
Qy 767 HisArgLeuSerAlaileGlnAsnAlaAspLeuileValValLeuHisAsnGlyLysIle 786
Db 3710 CACCGCTGTCCACCATCCAGAACCGCAGACTGATCGTGGTGATTCAGAACCGCCAGT 3769
Qy 787 LysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuVal 806
Db 3770 AAGGAGCAGCGGCCACCCAGCAGCGCTGCTGGCCCAAGAGGCATCTATTTCTCGATGGT 3829
Qy 807 AsnAla 808
Db 3830 CAGGCT 3835
RESULT 9
US-08-752-447-1
; Sequence 1, Application US/08752447
; Patent No. 5994088
; GENERAL INFORMATION:
; APPLICANT: Mechtner, Eugene
; TITLE OF INVENTION: Methods and Reagents for Preparing and
; TITLE OF INVENTION: Using Immunological Agents Specific for p-glycoprotein
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boehrnen Hulbert & Berghoff Ltd.
; STREET: 300 South Wacker Drive, Seventh Floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/752.447
; FILING DATE: 15-NOV-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5994088nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 95,1121
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-9808
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4669 base pairs

/ TYPE: nucleic acid
 / STRANDEDNESS: single
 / TOPOLOGY: linear
 / MOLECULE TYPE: cDNA
 / FEATURE:
 / NAME/KEY: 5'UTR
 / LOCATION: 1..424
 / FEATURE:
 / NAME/KEY: CDS
 / LOCATION: 425..4264
 / FEATURE:
 / NAME/KEY: 3'UTR
 / LOCATION: 4265..4669
 / US-08-752-447-1

Alignment Scores:

Pred. No.: 5,29e-272 Length: 4669
 Score: 2412.00 Matches: 462
 Percent Similarity: 76.06% Conservative: 164
 Best Local Similarity: 56.14% Mismatches: 183
 Query Match: 59.13% Indels: 14
 DB: 2 Gaps: 4

US-09-873-409-2 (1-812) x US-08-752-447-1 (1-4669)

QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
 DB 1781 GTTGATGGACAGCATATTAGGACCATAAATGTAAGGTTTCTACGGGAATCATTTGGTGTG 1840
 QY 22 ValSerGlnProValLeuPheGlyThrIleSerAsnAsnIleIysTyrGlyArg 41
 DB 1841 GTGAGTCAGGAACCTGTATTGTTTGGCCACCACCATAGCTGAAAACATTCGCTATGGCCGT 1900
 QY 42 AspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
 DB 1901 GAAATGTCCACATGATGATGATGAGATGAGAAAGCTGTCAAGGAACCAATGCTATGACTTT 1960
 QY 62 IleMetGluPheProAsnIysPheAsnThrLeuValGlyGluIysGlyAlaGlnMetSer 81
 DB 1961 ATCATGAACCTGCTCATAAATTTGACACCTCTGTTGGAGAGAGGGGCCAGTTGAGT 2020
 QY 82 GlyGlyGlnIysGlnArgIleAlaAlaArgAlaLeuValArgAsnProIysIleLeu 101
 DB 2021 GGTGGGAGAGAGAGGATCGCATTTGACGTGCGCTGTTGCAACCCCAAGATCCTC 2080
 QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerIysSerAlaValGlnAlaAla 121
 DB 2081 CTGCTGATGAGGCCAGCTGACGCTTGAGACAGAAAGCGAAGCAGTGGTTTCAGGTGGCT 2140
 QY 122 LeuGluIysAlaSerIysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
 DB 2141 CTGGATAAGGCCAGAAAAGTTCGACCACCATTTGTGATAGCTCATCTGTTTGTCTACAGTT 2200
 QY 142 ArgSerAlaAspLeuIleValThrLeuIysAspGlyMetLeuAlaGluIysGlyAlaHis 161
 DB 2201 CGTAATGCTCAGCTCATCGCTGTTTTCGATGATGGAGTCAATTTGGGAAGAAAGAAATCAT 2260
 QY 162 AlaGluMetAlaIysArgGlyLeuTyrTyrSerLeuValMetSerGln----- 178
 DB 2261 GATGAATCATGAAGAAGAAAGGCAATTTACTTCAAACTTGTCAATGTCAGACAGCAGGA 2320
 QY 179 -----AspIleIysIysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu 195
 DB 2321 AATGAAGTTGCAATAGAAATGACGCTGATGATGAAATCCAAAGTGAATTTGATGCTTGGAA 2380
 QY 196 ArgIysThrAsnSerLeuProLeuHisSerVal-----LysSerIle 209
 DB 2381 ATGCTCTCAAATGATTACAGATCCAGTCTCTAATAAGAAAAAAGATCAATCGTAGGAGTGC 2440
 QY 210 LysSerAspPheIleAspIysAlaGluGluSerThr--GlnSerIysGluIleSerLeu 228
 DB 2441 CGTGGATCACAGCCCNACACAGAAAGCTTAGTACCAAGAGGCTCTGGATGAAGATATA 2500

QY 229 ProGluValSerLeuLeuIysIleLeuIysLeuAsnLysProGluTyrProPheValVal 248
 DB 2501 CCTCCAGTTTCTCTTTGGAGGATTATGAAGCTAAATTTAACTCAATGGCTTATTGTT 2560
 QY 249 LeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePhe 268
 DB 2561 GTTGGTGTATTTTGGCCATTATAAATGGAGCCCTGCACACAGCATTTTCAATAATATT 2620
 QY 269 AlaIysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeuIysHisAspAla 287
 DB 2621 TCAGAATATTATAGGGGTTTTTACAGAATTGATGATCTCTGAAACAAACACAGATACT 2680
 QY 288 GluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMet 307
 DB 2681 AACTTGTCTTCTACTATTGTTCTAGCCCTTGGAAATATTATTTCTTATTACATTTTCTCT 2740
 QY 308 GlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla 327
 DB 2741 CAGGGTTTCACATTTGGCAAGCTGGAGAGATCCTCACCAAGCGGCTCGATACATGTT 2800
 QY 328 PheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGluLysGluAsnSerThrGly 347
 DB 2801 TTCCGATCCATGCTCAGACAGGATGTGAGTTGTTTCATGACCTCAAAACACCACTGGA 2860
 QY 348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg 367
 DB 2861 GCATTTGACTTACCAGGCTCGCCAAATGATGCTGCTCAAGTTAAAGGGGCTATAGGTTCCAG 2920
 QY 368 IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387
 DB 2921 CTTGCTGTAATTTACCAGAAATATAGCAATCTTGGACAGGAATTAATATATCTTCACTC 2980
 QY 388 TyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGly 407
 DB 2981 TATGTTGGCAACTTAACACTGTTACTCTTAGCAATTTGTACCATCATTCATAGCAGGA 3040
 QY 408 MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspIysGlnIleLysHis 427
 DB 3041 GTTGTGTAATGAAATGTTTCTCGACACGACACTGAAAGATAAGAAAGAACTAGAAAGT 3100
 QY 428 AlaGlyIysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArg 447
 DB 3101 GCTGGGAAGATCGCTACTGAAGCAATAGAAAACCTTCGGAACCGTTGTTCTTTGACTCAG 3160
 QY 448 GlyLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSer 467
 DB 3161 GAGCAGAACTTTGAACATATGATGCTCAGAGTTTGCAGGTACCATACAGAACTCTTTG 3220
 QY 468 LysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487
 DB 3221 AGGAAAGCACACATCTTTGGAATTTACATTTTCTTCCACCCAGGCAATGATGATTTTCTC 3280
 QY 488 TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507
 DB 3281 TATGCTGGAGTTTTCGGTTTGGAGCCTACTTGGTGGCAATATAAATCATGAGCTTTGAG 3340
 QY 508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyValaMetAlaIleGlyLysThrLeu 527
 DB 3341 GATGTTCTGTTAGTATTTTCAGCTGTTGTTGTCATGTCCTGGCGGCGNAGTCAGT 3400
 QY 528 ValLeuAlaProGluTyrSerIysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeu 547
 DB 3401 TCATTTGCTCTCTGACTATGCCAAAGCCAAATATCAGACAGCCACATCATGATCAT 3460
 QY 548 GluLysLysProAsnIleAspSerArgSerGlnGluIysIysProAspThrCysGlu 567
 DB 3461 GAAATAAACCCCTTTGATTGACAGCTACAGCAGGAAGGCTTAATGCCCGAACACATTGGAA 3520
 QY 568 GlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIle 587
 DB 3521 GGAAATGTACATTTTGGTGAAGTTGTATTCAACTATCCCAACCCGCGACATCCCGAGT 3580
 QY 588 LeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSer 607

QY 179 -----AspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu 195
Db 2321 AATGAAGTTGAATAGAAAATGACGCTGATGAATCAAAAGTCAAAATGATCGCTGGAA 2380
QY 196 ArgLysThrAsnSerLeuProLeuHisSerVal-----LysSerIle 209
Db 2381 ATGCTCTCAAGTGAATCAAGATCCAGCTCAATTAAGAAAAGATCAACTCGTAGGAGTGC 2440
QY 210 LysSerAspPheIleAspLysAlaGluGluSerThr---GlnSerLysGluIleSerLeu 228
Db 2441 CGTGGATCAACAGCCCAACACAGAAAGCTAGTACCAAGAGGCTCTGGATGAAGTATA 2500
QY 229 ProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValVal 248
Db 2501 CCTCCAGTTTCCTTTGGAGGATTATGAAGCTAAATTTAACTGAATGGCTTATTTTGTT 2560
QY 249 LeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePhe 268
Db 2561 GTTGGGTATTTTGTGCTCAATTAATTAATGGAGCTCTCAACCAAGCATTTTCAATATATTT 2620
QY 269 AlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeuLysHisAspAla 287
Db 2621 TCAGAGATTATAGGGGTTTTTACAGAAATTGATGATCTTGAAACAAACACAGAGATAGT 2680
QY 288 GluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMet 307
Db 2681 AACTCTGTTTCACTATTGTTTCTAGCCCTTGGAAATTATTTCTTTTATTACATTTTCTTT 2740
QY 308 GlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla 327
Db 2741 CAGGGTTTTCACATTGTGGCAAGCTGGAGAGATCTCTCACCAGCGCTCCGATACATGGTT 2800
QY 328 PheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGly 347
Db 2801 TTCCGATCAATGCTCAGACAGAGATGAGTGGTTTCTATGACCTTAAACACCACTGGA 2860
QY 348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg 367
Db 2861 GCATTGACTACCAAGCTCGCAATGATGCTGCTCAAGTTAAAGGGGCTATAGGTTCCAG 2920
QY 368 IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387
Db 2921 CTGCTGTAATTAACCCAGAAATATAGCAAAATCTTGGACAGGAATAATATATCTTCATC 2980
QY 388 TyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGly 407
Db 2981 TATGGTTGGCACTAACACTGTTACTTCTTAGCAATTTGTACCCATCATTCGAATAGCAGA 3040
QY 408 MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHis 427
Db 3041 GTTGTGTAATGAAATGTTTGTGTCGACAGCACTGAAAGATAAGAAAGAACTAGAAAGT 3100
QY 428 AlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArg 447
Db 3101 GCTGGGAAGATCGCTACTGAAGCAATAGAAAACCTCCGAACCGCTTCTTTTGTGACTCG 3160
QY 448 GluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSer 467
Db 3161 GAGCAAGATTGAAACATATGATGCTCAGAGTTTGAGGTACCATACAGAAACTCTTTG 3220
QY 468 LysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487
Db 3221 AGGAAGCACACATCTTTGGAATTACATTTCTTCCACCAGGCAATGATGATTTTTC 3280
QY 488 TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507
Db 3281 TATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACAATAAATCTCATGAGCTTTGAG 3340
QY 508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeu 527
Db 3341 GATGTTCTGTATGATTTTTCAGCTGTGTCTTTGGTGGCCATGGCCGCTGGGGCAAGTCAGT 3400

QY 528 ValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeu 547
Db 3401 TCATTTGCTCTGACTATGTCAAAAGCCAAAATATATCAGCAGCCCATCATCATGATGATTT 3460
QY 548 GluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGlu 567
Db 3461 GAAAAAACCCCTTTGATTGACAGCTACAGCAGCGAAGGCTTAATCCCGACACATTTGAA 3520
QY 568 GlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIle 587
Db 3521 GGAAATGTCACATTTGGTGAAGTTGATTTCAACTATCCACCCGACCGACATCCAGTG 3580
QY 588 LeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSer 607
Db 3581 CTTTCAGGACTGAGGCTGGAGTGAAGAGGCGCAGACGCTGGCTCTGGTGGCCAGCAGT 3640
QY 608 GlyCysGlyLysSerThrSerValGlnLeuGlnArgLeuTyrAspProValGlnGly 627
Db 3641 GGCTGTGGAGAGACAGCTGGTCCAGCTCTGGAGCGGTTCTACGACCCCTTGGCAGGG 3700
QY 628 GlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGln 647
Db 3701 AAGTGTCTTGTGATGGCAAGAAATAAAGCGACTGAATGTTCACTGGCTCCGACGACAC 3760
QY 648 IleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAla 667
Db 3761 CTGGGCTGCTGCTCCAGGAGCCCATCTCTGTTGATGCTCAGCATTTGCTGAGAACATTC 3820
QY 668 TyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAlaAla 687
Db 3821 TATGAGACAAACGCGGGTGTGTCAGGAAGAGATCTGTCAGGCGCAGCAAGAGGCC 3880
QY 688 AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLys 707
Db 3881 AACATACATGCTTCTCATCGAGTCACTGCTTAATAAATATAGCACTAAAGTAGCAGACAA 3940
QY 708 GlyAlaGlnLeuSerGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuGln 727
Db 3941 GGAACCTCAGCTCTCTGGTGGCCAGAAACAAACGCAATGGCATAGCTGCTGCTTGTAGA 4000
QY 728 LysProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 747
Db 4001 CAGCTCATATTTCTTCTTGGATGAGCCAGCTCAGCTCTGGATACAGAAAGTGAAG 4060
QY 748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767
Db 4061 GTTGTCCAAAGAGCCCTGGACAAAGCCAGAGAGGCGCACCTGCATTGTGATTGCTCAC 4120
QY 768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys 787
Db 4121 CGCCTGTCCACCATCCAGAAATGCAACTTAATAGTGGTGTTCAGAAATGGCAGAGTCAAG 4180
QY 788 GluGlnGlyThrHisGlnLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807
Db 4181 GAGCATGGCAGCATCAGCAGCTGCTGGCAGAAAGGCACTATTTTCAATGTCAGT 4240
QY 808 AlaGlnSer 810
Db 4241 GTCCAGCT 4249
RESULT 11
US-08-583-276-18
; Sequence 18, Application US/08583276
; Patent No. 5837536
; GENERAL INFORMATION:
; APPLICANT: McDonagh, Kevin T.
; APPLICANT: Nienhuis, Arthur
; APPLICANT: Tolstoshev, Paul
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HUMAN
; TITLE OF INVENTION: MULTIDRUG RESISTANCE GENES AND IMPROVED
; TITLE OF INVENTION: SELECTION OF CELLS TRANSDUCED WITH SUCH GENES
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:

ADDRESS: Carella, Byrne, Bain, Gilfillan,
ADDRESS: Cecchi & Stewart
STREET: 6 Becker Farm Road
CITY: Roseland
STATE: New Jersey
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch diskette
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: DW4.V2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/583.276
FILING DATE: 05-JAN-1996
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/332,444
FILING DATE: 31-OCT-1994
APPLICATION NUMBER: 07/887,712
FILING DATE: 22-MAY-1992
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 4669 bases
TYPE: nucleic acid
STRANDEDNESS: singular
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
US-08-583-276-18

Alignment Scores:

Pred. No.: 3,046-270 Length: 4669
Score: 2397.00 Matches: 460
Percent Similarity: 76.03% Conservative: 165
Best Local Similarity: 55.96% Mismatches: 183
Query Match: 58.76% Indels: 14
DB: 2 Gaps: 4

US-09-873-409-2 (1-812) x US-08-583-276-18 (1-4669)

Qy 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyArgAspHisIleGlyVal 21
Db 1781 GTTCATGGACAGATATTAGGACCAATAATGTAAGGTTTCTACGGGAATCATTTGGTGTG 1840
Qy 22 ValSerGlnGluProValLeuPheGlyThrIleSerAsnAsnIleLysTyArgGly 41
Db 1841 GTGAGTCAGGAACCTGTATTGTTGCCACACGATAGCTGAAACACATTTCGCTATGGCCGT 1900
Qy 42 AspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyArgPhe 61
Db 1901 GAAATGTCCATCATGATGAGATTGAGAAAGCTGTCAAGNAGCCATGCTATGACTTT 1960
Qy 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
Db 1961 ATCATGAAATCGCTCATAAATTTGACCCCTGGTTGGAGAGAGGGCCCGCCAGTTGAGT 2020
Qy 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
Db 2021 GTGGGCGAAGCAGAGGATCGCCATTGTCAGCTGCCCTGGTTCCGAAACCCCAAGATCCCTC 2080
Qy 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
Db 2081 CTGCTGATGAGCCACGTCAGCTTGGACACAGAAAGCGAGCAGTGGTTTCAGGTGGCT 2140
Qy 122 LeuGluLysAlaSerLysGlyArgThrIleValValAlaHisArgLeuSerThrIle 141
Db 2141 CTGATAAGCCAGAAAGGTTCGACCAACCATTTGTATAGCTCATCGTTTGTCTACAGTT 2200
Qy 142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
Db 2201 CGTAATGCTGACGTCACTCGTGTTCGATGATGGAGTCAATTTGTGGAGAAAGGAATCAT 2260

Qy 162 AlaGluLeuMetAlaLysArgGlyLeuTyThrSerLeuValMetSerGln----- 178
Db 2261 GATGAACATCATGAAAGAGAAAGGCAATTTACTTCAAACTTGTCAACATGACAGACAGGA 2320
Qy 179 -----AspIleLysLysAlaAspGluGlnMetGluSerMetThrThrTyThrGlu 195
Db 2321 AATGAAGTTGAAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAATGATGCTTCTTGGAA 2380
Qy 196 ArgLysThrAsnSerLeuProLeuHisSerVal-----LysSerIle 209
Db 2381 ATGTCTTCAAAATGATTCAAGATCACTTAATAAGAAAAAGATCAACTCGTAGGAGTGTC 2440
Qy 210 Lys-----SerAspPheIleAspLysAlaGluGluSerThrGlnSerLysGluLysSerLeu 228
Db 2441 CGTGATCATCAACCCCAAGAGCTTAGTACAAAGAGGCTCTGGATGAAAGTATA 2500
Qy 229 ProGluValSerLeuLysIleLeuLysLeuAsnLysProGluTyProPheValVal 248
Db 2501 CCTCAGTTTCTTTGGAGGATTATGAAGCTAAATTTAACTGAATGGCCTTATTTTGT 2560
Qy 249 LeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePhe 268
Db 2561 GTTGGTGTATTTTGTGCCATTATAAATGGAGCCTGCAACAGCAGCATTTGCAATAATATT 2620
Qy 269 AlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeuLysHisAspAla 287
Db 2621 TCAAGATTTATAGGGGTTTTTCAAGAATTTGATGATCTCTGAAACAAACGACAGAACTACT 2680
Qy 288 GluIleTyThrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyThrPheMet 307
Db 2681 AACTTGTTCACATATGTTCTAGCCCTTGGAAATTTCTTTTATTACATTTTTCCT 2740
Qy 308 GlnGlyLeuPheTyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla 327
Db 2741 CAGGGTTTCACATTTGGCAAGCTGGAGAGATCTCACCAGCGGCTCGGATACATGTT 2800
Qy 328 PheLysAlaMetLeuTyArgGlnAspIleAlaTyPheAspGluLysGluAsnSerThrGly 347
Db 2801 TTCCGATCCATGCTCAGACAGATGTGAGTTGTTGTATGACCTTAAACACACCACTGGA 2860
Qy 348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg 367
Db 2861 GCATGTGACTACAGGCTCGCCAAATGATGCTGCTCAAGTTAAAGGGGCTATAGTTCACAG 2920
Qy 368 IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387
Db 2921 CTGCTGTAAATTCACAGATATAGCAATTCCTGGACAGGAAATATATATCTTCTTCATC 2980
Qy 388 TyrGlyTyPgluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGly 407
Db 2981 TATGTTGGCAACTAACACTGTTACTCTTAGCAATTTGACCCATCATTCGAATAGCAGGA 3040
Qy 408 MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLysHis 427
Db 3041 GTTGTGAAATGAAATGTTGTCTGGCAAGCAGCTGAAAGATAAGAAAGAACTAGAAAGT 3100
Qy 428 AlaGlyLysIleAlaThrGluAlaLeuAsnIleArgThrIleValSerLeuThrArg 447
Db 3101 GTTGGAGATGCTGCTACTGAGCAATAGAAAATTCCTCGAACCGTGTCTTCTTGACTCAG 3160
Qy 448 GluLysAlaPheGluGlnMetTyThrGluMetLeuGlnThrGlnHisArgAsnThrSer 467
Db 3161 GAGCAGAAAGTTGAACATATGATGCTCAGAGTTTGCAGTGTGACCTACCATCAGAAACTCTTT 3220
Qy 468 LysLysAlaGlnIleIleGlySerCysTyThrAlaPheSerHisAlaPheIleTyThrPheAla 487
Db 3221 AGGAAAGCAGACATCTTGGAAATTTACATTTCTTCCACCGGCAATGATGATTTTCTCC 3280
Qy 488 TyrAlaAlaGlyPheArgPheGlyAlaTyThrLeuIleGlnAlaGlyArgMetThrProGlu 507
Db 3281 TATGCTGGATGTTTCGGTTTGGAGCTTCTTGGTGGCACATAAACTAATGAGCTTTGAG 3340
Qy 508 GlyMetPheIleValPheThrAlaIleAlaTyThrGlyAlaMetAlaIleGlyLysThrLeu 527


```
Db 3341 GATGTTCTGTAGTATTTCAGCTGTGTCTTTGGTGGCCATGCGCTGGGCAAGTGAGT 3400
Qy 528 ValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeu 547
Db 3401 TCATTTGCTCCTGATATCCCAAGCAAAATATACAGAGCCACATCATCATGATCAT 3460
Qy 548 GluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGlu 567
Db 3461 GAAAAAACCCCTTGTATTGACAGTACAGCAGGAAGGCTTAATGCCGACACATTTGGAA 3520
Qy 568 GlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIle 587
Db 3521 GGAATGTCACATTTGCTGAAGTTGTATTCACTATCCACCCGACCGACATCCCAAGTC 3580
Qy 588 LeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSer 607
Db 3581 CTTTCAGGACTGAGCTGGAGGTGAAGAAGGCGCACACGCTGGCTCTGGTGGGCACAGT 3640
Qy 608 GlyCysGlyLysSerThrSerValGlnLeuGlnArgLeuTyrAspProValGlnGly 627
Db 3641 GGCCTGGGAAGACACAGTGGTCCAGCTCCTGGAGGGGTTCTACGACCCCTTGGCAGGG 3700
Qy 628 GlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGln 647
Db 3701 AAGTGTCTGCTCATGGCAAGAAATAAAGCGACTGAATGTTCAAGTGGCTCCGAGCACAC 3760
Qy 648 IleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAla 667
Db 3761 CTGGGATCGTGTCCAGGAGCCCTCTGTTGAGTCGACGATTCGTGAGAACATTTGCC 3820
Qy 668 TyrGlyAspAsnSerArgValValProLeuAspGluLysGluAlaAlaAsnAlaAla 687
Db 3821 TATGGACACACAGCGGGTGGTGTACAGGAAGAGATCTGAGGGCAGCAAGAGGAGGCC 3880
Qy 688 AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLys 707
Db 3881 AACATACATGCTTCATCGAGTCACTGCCTTAATAATATAGCACTAAAGTAGGAGACAAA 3940
Qy 708 GlyAlaGlnLeuSerGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGln 727
Db 3941 GGAACCTCAGCTCTCTGTGGCCAGAACACACGCAATTCGCCATGCTGTGCTGTAGA 4000
Qy 728 LysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 747
Db 4001 CAGCCTCATATTTGCTTTGGATGAAGCCACGTCAGCTCTGGATACAGAAAGTGAAG 4060
Qy 748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767
Db 4061 GTTGTCCAGAAAGCCCTGGACAAACCCAGAGAGGCGCACCTGCATGTTGTTGCTCAC 4120
Qy 768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisGlnGlyLysIleLys 787
Db 4121 CGCCTGTCCACCATCCAGATGACAGATTAATAGTGTGTTTCAGAAATGGCAGAGTCAAG 4180
Qy 788 GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807
Db 4181 CAGATGGCAGCATCAGCAGTCTGCTGGCACAGAAAGGCAATCTATTTTCAATGGTCAGT 4240
Qy 808 AlaGln 809
Db 4241 GTCCAG 4246
```

RESULT 12

```
US-08-461-823-1
; Sequence 1, Application US/08461823
; Patent No. 5593840
; GENERAL INFORMATION:
; APPLICANT: Bhatnagar, Satish K.
; APPLICANT: George Jr., Albert L.
; APPLICANT: Nazarenko, Irina
; TITLE OF INVENTION: AMPLIFICATION OF NUCLEIC ACID SEQUENCES
; NUMBER OF SEQUENCES: 27
```

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OncorPharm, Inc.
; STREET: 200 Perry Parkway
; CITY: Gaithersburg
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20877
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,823
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/168,621
; FILING DATE: 16-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/010,433
; FILING DATE: 27-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Karta, Glenn E.
; REGISTRATION NUMBER: 30,649
; REFERENCE/DOCKET NUMBER: PA-0012 CIP 2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301 527-2058
; TELEFAX: 301 208-6997
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2726 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-461-823-1

Alignment Scores:
Pred. No.: 6,35e-257 Length: 2726
Score: 2280.00 Matches: 435
Percent Similarity: 76.55% Conservative: 159
Best Local Similarity: 56.06% Mismatches: 168
Query Match: 55.90% Indels: 14
DB: 1 Gaps: 4

US-09-873-409-2 (1-812) x US-08-461-823-1 (1-2726)
Qy 49 MetGluArgAlaAlaArgGluAlaAlaAlaTyrAspPheIleMetGluPheProAsnLys 68
Db 2 ATTGAGAAAGCTGTCAAGGAAGCCAAATGCTATGACTTTATCATGAAGCTGCTCATAAA 61
Qy 69 PheAsnThrLeuValGlyGlyLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIle 89
Db 62 TTTGACACCTGTTGGACAGAGAGGGGCCAGTTGAGTGTGGGCGAAGCAGAGGATC 121
Qy 89 AlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleuAspGluAlaThrSer 108
Db 122 GCCATTGCAGCTGCCCTGTTGCCAACCCCAAGATCTCTCTGCTGGATGAGGCCAGTCA 181
Qy 109 AlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGly 128
Db 182 GCCTTGGACACAGAAAGCCAGAGCTGGTTCAGGTGGCTCTGGATGAAGCCAGAAAGGT 241
Qy 129 ArgThrThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleVal 148
Db 242 CGGACCAACCATTTGTAGTATCGTTTGTCTACAGTTGCTGAATGCTGACGTCATCGCT 301
Qy 149 ThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArg 168
Db 302 GGTTCGATGATGGAGTCATTGTGGAGAAAGGAATCATGATGAATCATGAAGAGAAA 361
```


Qy	515	AlaIleAlaTyrGlyAlaMetAlalleGlyysThrLeuValLeuAlaProGluTyrSer	534
Db	1442	GCTGTGTCCTTTGGTGCATGCGCGGGGCAAGTCAGTTCATTGTCCTCGATGATATGCC	1501
Qy	535	LysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAsp	554
Db	1502	AAAGCCAAATATCAGCAGCCCATCATGATCATTTGAAAAAACCCCTTTGATTGAC	1561
Qy	555	SerArgSerGlnGluGlyLysProAspThrCysGluGlyAsnLeuGluPheArgGlu	574
Db	1562	AGCTACAGCAGCGAAGCGCTAATATGCGGAACATATTGGGAAGGAAATGTCCACATTTGGTGAA	1621
Qy	575	ValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSer	594
Db	1622	GTGTGTTTCAACTATCCACCCGACCGAGATCCCAAGTGTCTTCAGGACTCAGCCTGGAG	1681
Qy	595	IleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSer	614
Db	1682	GTGAAGAAGGCCAGACGCTGCTCTGTGTGGCAGCAGTGGCTGTGGGAAGACACACAGTG	1741
Qy	615	ValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyVal	634
Db	1742	GTCAGCTCTCGAGCGGTCTTACGACCCCTTGGCAGGGAAGTGTCTCTTGATGGCAAA	1801
Qy	635	AspAlaLysGluLeuAsnValGlnTyrLeuArgSerGlnIleAlaIleValProGlnGlu	654
Db	1802	GAATAAAGCGACTGAATGTTTCAGTGGCTCCGAGCAGCACCTGGGGCATCGTCTCCAGGAG	1861
Qy	655	ProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgVal	674
Db	1862	CCCATCTGTTTGACTGCGACATTGCTGAGAACATTGCTATGAGACACACAGCGGGTG	1921
Qy	675	ValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGlu	694
Db	1922	GTGTCAAGGAAGAGATCGTGGAGGCGAGCAAAAGGAGGCCAACATACATGCTTCATCGAG	1981
Qy	695	GlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGly	714
Db	1982	TCAGTGCCTAATAATATAGCACTAAAGTAGAGACAAAGGAACTCAGCTCTCTGGTGGC	2041
Qy	715	GlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeu	734
Db	2042	CAGAAACACGCATTCGCATAGCTGTGCGCTTGTATAGACGCTCATATTTGCTTTTG	2101
Qy	735	AspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAsp	754
Db	2102	GATGAAGCCACGTCAGCTCTGGATACAGAAAGTGAAAGTTGTCCAAAGAGCCCTGGAC	2161
Qy	755	LysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsn	774
Db	2162	AAAGCCAGAGAGCGCGACCTGCATTTGATTTGTCACCGCTGTCCACCATCCAGAAAT	2221
Qy	775	AlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlu	794
Db	2222	GCAGACTTAATAGTGTGTTTCAGAATGCGCAGATCAAGGAGCATGGCACCATCCAGCAG	2281
Qy	795	LeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer	810
Db	2282	CTGTGGCACAGAAAGCATCTATTTTCAATGGTTCAGTGTCCAGGCT	2329
RESULT 13			
US-08-996-545-1			
; Sequence 1, Application US/08996545			
; Patent No. 5928898			
; GENERAL INFORMATION:			
; APPLICANT: Skatrud, Paul L.			
; APPLICANT: de Waard, Maarten A.			
; APPLICANT: Peery, Robert B.			
; APPLICANT: Andrade, Alan C.			
; TITLE OF INVENTION: Multiple Drug Resistance Gene at rD of			
; TITLE OF INVENTION: Aspergillus nidulans			
; NUMBER OF SEQUENCES: 3			

; NUMBER OF SEQUENCES: 3

; NUMBER OF SEQUENCES: 3

ADDRESSEE: Eli Lilly and Company
 STREET: Lilly Corporate Center
 CITY: Indianapolis
 STATE: Indiana
 COUNTRY: U.S.
 ZIP: 46285
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/996,545
 FILING DATE:
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Webster, Thomas D.
 REGISTRATION NUMBER: 39,872
 REFERENCE/DOCKET NUMBER: X-11766
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 317-276-3334
 TELEFAX: 317-276-2763
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 4002 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..4002
 US-08-996-545-1

Alignment Scores:

Pred. No.: 1 87e-187 Length: 4002
 Score: 1690.50 Matches: 366
 Percent Similarity: 61.52% Conservative: 152
 Best Local Similarity: 43.47% Mismatches: 293
 Query Match: 41.44% Indels: 31
 DB: 2 Gaps: 7

US-09-873-409-2 (1-812) x US-08-996-545-1 (1-4002)

Qy 1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
 Db 1465 TTGCTGGATGGCCATGATGATCAAGGACCTCAATCTCCGCTGGCTTCGCCAACAGATCTCT 1524
 Qy 21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAnilleLysTyrGly 40
 Db 1525 TTGCTTAGCAGGAGCGCTGTTCTTTTGGCAGCAGATTTATAGAATATTAGGCGCGGT 1584
 Qy 41 -----ArgAspAspValThrAspGluGluMetGluArg 51
 Db 1585 CTCATCCGCACAAAGTACGAGATGAATCCGAGGATAAGTCCGGGAACTCATCGAGAAC 1644
 Qy 52 AlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThr 71
 Db 1645 GCGGCAAAATGGCGAATGCTCATGACTTTTATTACTGCTTGCCTGAAGGTTATAGGCGAG 1704
 Qy 72 LeuValGlyGluLysGlyAlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAla 91
 Db 1705 AATGTTGGGCGAGCGTGGCTTCTCTCTTTCAGGTGGCCAGAACAGCGCATTGAATCGCC 1764
 Qy 92 ArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAsp 111
 Db 1765 CGTGGCGGTGTAGTGACCCCAAAATCTGCTCTGATGAGGACTCTCGGCGCTTGAC 1824
 Qy 112 SerGluSerLysSerAlaValGlnAlaLeuGluLysAlaSerLysGlyArgThrThr 131
 Db 1825 ACAAATCCGAAGGCGTGGTTCAAGCAGCTTTGGAGAGGGCAGCTGAAGGCCGAATCTACT 1884

Qy 132 IleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLys 151
 Db 1885 ATTGTGATCGCTCATCGCTTTCCACGATCAAAACGGCGCAACACATTGTGGTCTGGTC 1944
 Qy 152 AspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyr 171
 Db 1945 AATGGCAAAATTGCTGAACAAGAACTCAGATGAATTTGTTGACCGCGAGCGCTTAT 2004
 Qy 172 TyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThr 191
 Db 2005 CGCAAACTTGTGGAGGCTCAACGATCAATGAACACAAAGAACTGACGCTTGGAGGAC 2064
 Qy 192 TyrSerThrGluArgLysThrAsnSer-----LeuProLeuHisSerValLysSerIle 209
 Db 2065 GCGACGCTGAGGATCTCAGGAATGCAGATATTTGCCAAATCAAAACTGCGTCAAGCGCA 2124
 Qy 210 LysSerAspPhe-----IleAspLys-----AlaGluGluSerThr 221
 Db 2125 TCATCGGATCTCGACGGAACCCACACCATTTGACCGCAGCGGACCCACACATCTGTT 2184
 Qy 222 GlnSerLysGluIleSer-----LeuProGluValSerLeuLeuLys 235
 Db 2185 TCACGCGGATTTCTTCTAAAGACCCCGGAAACCAACTCCGAAATACTCATTTATGACG 2244
 Qy 236 IleLeuLys-----LeuAsnLysProGluTrpProPheValValLeuGlyThr 251
 Db 2245 CTGCTCAAAATTTGTTGCTTCTTCAACCGCCCTGAAATCCGTTACATGCTCATCGGTCT 2304
 Qy 252 LeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIle 271
 Db 2305 GTCTTCTCAGTGTAGTGTGGTGGCCAAACCCACCAAGCAGTGTATATGATAAGCC 2364
 Qy 272 IleThrMetPheGly-----AsnAsnAspLysThrThrLeuLysHisAspAlaGluIle 289
 Db 2365 ATCAGCACACTCTCGCTCCAGAAATCACAATATAGCAAGCTTCGACATGCGGATTTTC 2424
 Qy 290 TyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGly 309
 Db 2425 TGCTCATTTGATGTTCTTCGTTGTTGTTATCATTCAGTTTATCAGCGCAGTCAACCAATGGT 2484
 Qy 310 LeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLys 329
 Db 2485 GCTGCATTTGCGGTATGCTCCGAGAGACTTATTGCTCGCGCAGAGACACTGCTCTTCGG 2544
 Qy 330 AlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeu 349
 Db 2545 ACATATCTCGTCAAGACATTTGCTTTTTCAGCAAGAGAGAAATAGCCCGCGCTCTG 2604
 Qy 350 ThrThrIleAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGly 369
 Db 2605 ACCTCTTCTCTCCACCGAGACGAGCATCTCTCCGGTGTAGCGGTGTGACTCTAGGC 2664
 Qy 370 ValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGly 389
 Db 2665 ACGATCTTGATGACCTCCACGACCTTAGGAGCGGCTATCATTTATGCGGTGGGATTGGG 2724
 Qy 390 TrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIle 409
 Db 2725 TGAATATGGCTTGGTTGATCTCGGTGTGCGGTTCTCTCTGCGATCGGTTTCTTAC 2784
 Qy 410 GluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLysHisAlaGly 429
 Db 2785 CGATTTCTATGCTAGCCCACTTTCATCAGCTCCAGCTCCAGGCTGCTTATGAGGATCTGCA 2844
 Qy 430 LysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLys 449
 Db 2845 AACTTTGCTTGGAGGCTACATCGTCTATCCGACAGTTCGCTCATTAACCCGGAAGG 2904
 Qy 450 AlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLys 469
 Db 2905 GATGCTGGGAGATTTTACCATGCCAGCTTTCAGCGCAAGGCGAGGACCTCTTAATCTCT 2964


```
QY 470 AlaGlnIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAla 489
DB 2965 GTCTTGAGGTATCCCTGTTATATGCGTCGTCGAGGCACCTGTTTCTCTCGGTGCG 3024
QY 490 AlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMet 509
DB 3025 CTCGGGTTTGTGACGAGGACACCTCTTGTGTCACCATCGCGGACCGCTCTTTCCCTTC 3084
QY 510 PheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeu 529
DB 3085 TTTGTTGTTCTCCGAGATTCTCTTGTGTCATCCGCGGACCGCTCTTTCCCTTT 3144
QY 530 AlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuGluLys 549
DB 3145 GCACCAAGATGGCAAGCGAAGATGCGCGCGCGCGAATTCGACGACTGTTCCGACGA 3204
QY 550 LysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsn 569
DB 3205 AAGCCAAATTTGATNACTGTTCTGAAGAGGCGGAGAGCTCGAAACGGTGAAGTGAA 3264
QY 570 LeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArg 589
DB 3265 ATCGAATTTAGGAACGTGCATTCAGATACCCGACCGCCAGACAGCTGTCCTGCGC 3324
QY 590 GlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCys 609
DB 3325 GCGTTGACCTGACCGTGAAGCTGACAAATATGTTGCGCTTGTGCGACCCGCGGTTGT 3384
QY 610 GlyLysSerThrSerValGlnLeuGlnArgLeuTyrAspProValGlnGlyVal 629
DB 3385 GGCAAGAGTACCACCATTCGATTGCTGTGAGCGCTTTTACGATCGGATGCCGGTCCATC 3444
QY 630 LeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAla 649
DB 3445 CTGTGTGATGGGAAGACATAGTAACCTAATATCACTCTACCGCAGCTTCTGTGCA 3504
QY 650 IleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGly 669
DB 3505 CTGGTCAGCGAGCGGACACATGTACACGAGGACCATCAAGGMAAACATCTTACTTGGT 3564
QY 670 AspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaIle 689
DB 3565 ATGTGCAAGATGACGTACCGGAAGAAATCTTGATTAAAGGCTTGCAAGGACGCTAATATC 3624
QY 690 HisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAla 709
DB 3625 TAGCACTTCATCATGCTCGCTCCGAGGCGCTTTAATACAGTTGTTGGCAGCAGGAGGC 3684
QY 710 GlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuGlnLysPro 729
DB 3685 ATGTTGCTGCGGCGCCAAAGCAACGTGTGGCCATTGCGCGAGCCCTTCTCGGATCCC 3744
QY 730 LysIleLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValVal 749
DB 3745 AAAATCTTCTTCTCGATGAAGCGACGTGAGCCCTCGACTCCGAGTCAGAAAGGTCGTC 3804
QY 750 GlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeu 769
DB 3805 CAGCGCGCTTTGGATCCCGTCGCGAGCGGACCAATCGCCGTTGCACACCGACTC 3864
QY 770 SerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGln 789
DB 3865 AGCAGATTCAAAGCGGAGGTTATCTATGTTTTCGACCAAGGCAAGATCGTCGAAGC 3924
QY 790 GlyThrHisGlnLeuLeuLeuArgAsnArgPheIleTyrPheLysLeuValAsnAlaGln 809
DB 3925 GGAACGCACAGCGAAGTGTCCAGAAAAAGGCGCGGTACTACGAGCTGTCTCACTTTCGAG 3984
QY 810 SerVal 811
DB 3985 AGCTTG 3990
RESULT 14
```

```
US-08-996-545-3
; Sequence 3, Application US/08996545
; Patent No. 5928898
; GENERAL INFORMATION:
; APPLICANT: Skatrud, Paul L.
; APPLICANT: de Waard, Maarten A.
; APPLICANT: Peery, Robert B.
; APPLICANT: Andrade, Alan C.
; TITLE OF INVENTION: Multiple Drug Resistance Gene atrd of
; TITLE OF INVENTION: Aspergillus nidulans
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: U.S.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/08/996,545
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Webster, Thomas D.
; REGISTRATION NUMBER: 39,872
; REFERENCE/DOCKET NUMBER: X-11766
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3334
; TELEFAX: 317-276-2763
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4002 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: mRNA
; HYPOTHEetical: NO
; ANTI-SENSE: NO
US-08-996-545-3

Alignment Scores:
Pred. No.: 1,87e-187 Length: 4002
Score: 1690.50 Matches: 366
Percent Similarity: 61.52% Conservative: 152
Best Local Similarity: 43.47% Mismatches: 293
Query Match: 41.44% Indels: 31
DB: 2 Gaps: 7

US-09-873-409-2 (1-812) x US-08-996-545-3 (1-4002)
QY 1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
DB 1465 UUGUGGAGGCGCAUGCAUCAAGACCAUCCUCCGUGGCUUCCGCCAACAGAUUCU 1524
QY 21 ValValSerGlnGluProValLeuPheGlyThrIleSerAsnAsnIleLysTyrGly 40
DB 1525 UUGGUUAGCCAGGAGCCUUCUUUUGGACGAGCAUUAUAAGAAUUAUAGGCACGGU 1584
QY 41 -----ArgAspAspValThrAspGluGluMetGluArg 51
DB 1585 CUCAUCCGCAAAAGUACGAGAAUCCGAGGAUAAAGUCCCGGGAACUACUACGAGAAC 1644
QY 52 AlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThr 71
DB 1645 GCGGCAAAAAGCGCAUUGCUCAUGACUUUAUACUUGCCUUGAGGUAGAGACC 1704
QY 72 LeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAla 91
```


QY 790 GlyThrHisGlnLeuLeuArgAsnArgAspIleTyrPhelYsLeuValAsnAlaGln 809
Db 3925 GGAACAGCAGCGAACUGGUCAGAAAGAGGCGCGUACUACGAGCUGGUCACUUGCAG 3984
QY 810 SerVal 811
Db 3985 AGCUUG 3990

RESULT 15

US-09-328-320-1
; Sequence 1, Application US/09328320
; Patent No. 6228615
; GENERAL INFORMATION:
; APPLICANT: Skatrud, Paul L.
; APPLICANT: de Waard, Maarten A.
; APPLICANT: Peery, Robert B.
; APPLICANT: Andrade, Alan C.
; TITLE OF INVENTION: Multiple Drug Resistance Gene at/d of
; TITLE OF INVENTION: Aspergillus nidulans
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: U.S.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/328,320
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/996,545
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Webster, Thomas D.
; REGISTRATION NUMBER: 39,872
; REFERENCE/DOCKET NUMBER: X-11766
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3334
; TELEFAX: 317-276-3763
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4002 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..4002

US-09-328-320-1

Alignment Scores:
Pred. No.: 1.87e-187 Length: 4002
Score: 1690.50 Matches: 366
Percent Similarity: 61.52% Conservative: 152
Best Local Similarity: 43.47% Mismatches: 293
Query Match: 41.44% Indels: 31
DB: 4 Gaps: 7

US-09-873-409-2 (1-812) x US-09-328-320-1 (1-4002)

QY 1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
Db 1465 TTGCTGGATGGCCATGACATCAAGGACCTCAATCTCGCTGGCTTCGCCAACAGATCTCT 1524

QY 21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleYsTyrGly 40
Db 1525 TTGGTTAGCCAGGAGCCTGTTCTTTTGGCAGCAGCATTTATAAGAAATATTAGGACGGT 1584
QY 41 -----ArgAspAspValThrAspGluMetGluAsg 51
Db 1585 CTCATCGGCACAAAGTACGAGAAATGAATCCGAGGATAAGTCGGGAACATCATCGAGAAC 1644
QY 52 AlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnIlySsPheAnThr 71
Db 1645 GCGGCAAAAATGCGAATGCTCATGACTTTTATTACTGCTTTCCTGCTGAAGTTATGAGACC 1704
QY 72 LeuValGlyGluYsGlyAlaGlnMetSerGlyGlyGlnIlySsGlnArgIleAlaIleAla 91
Db 1705 AATGTTGGCAGCGTGGCTTCTCTCTTCAGGTGGCCAGAAAACAGCGCATTCGAATCGCC 1764
QY 92 ArgAlaLeuValArgAsnProIlyValIleLeuIleLeuAspGluAlaThrSerAlaLeuAsg 111
Db 1765 CGTGGCTTGTAGTAGCAACCAAAATCTCTCTCGATGAAGTACTTCGGCTTGGAC 1824
QY 112 SerGluSerLysSerAlaValGlnAlaLeuGluLysAlaSerIlySsGlyArgThrThr 131
Db 1825 ACAAATCCGAAGCGTGGTTCAGCAGCTTTGGAGAGGCGAGCTGAAGGCCGAATCTACT 1884
QY 132 IleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLys 151
Db 1885 ATTGTGATCGCTCATCGCTTTCACGATCAAAACGGCGCAACAATTTGTGTCTTCTGGTC 1944
QY 152 AspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyr 171
Db 1945 AATGGCAAAATGCTGAACAAGGAATCAGATGAATTTGTTGACCGGAGCGCTTAT 2004
QY 172 TyrSerLeuValMetSerGlnAspIleYsLysAlaAspGluGlnMetGluSerMetThr 191
Db 2005 CGCAAACTTGTGGAGCTCAACGTATCAATGAACAGAAGAGCTCACGCTTGGAGGAC 2064
QY 192 TyrSerThrGluArgLysThrAsnSer-----LeuProLeuHisSerValIlySsSerIle 209
Db 2065 GCCGAGCTGAGGATCTCAGAAATGCAATATTTGCCAAATCAAAACTCGCTCAAGCGCA 2124
QY 210 LysSerAspPhe-----IleAspLys-----AlaGluGluSerThr 221
Db 2125 TCATCGATCTCAGCGGAAACCCACCAACATTCAGCCGCGGACCCCAAGTCTGTT 2184
QY 222 GlnSerLysGluIleSer-----LeuProGluValSerLeuLeuLys 235
Db 2185 TCCAGCGCGATTCTTTCTAAAGAGACCCCGGAAACAACTCCGAAATACTCATTTATGGACG 2244
QY 236 IleLeuLys-----LeuAsnLysProGluTyrProPheValValLeuGlyThr 251
Db 2245 CTGCTCAAAATTTGTTGCTTCTTCAACCGCTGAAATCCCGTACATGCTCATCGCTCT 2304
QY 252 LeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaIlySsIle 271
Db 2305 GTCTTCTAGTGTAGTGTGGTGGCCAAACCCCAACCAAGTGTATATGCTAAAGCC 2364
QY 272 IleThrMetPheGly-----AsnAsnAspLysThrThrLeuLysHisAspAlaGluIle 289
Db 2365 ATCAGCACACTCTCGCTCCAGAAATCACAATATAGCAAGCTTCGACATGATGGGATTTC 2424
QY 290 TyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGly 309
Db 2425 TGCTCATTTGATGTTCTTCGTGGTGTGTTATCATTCAGTTTATCAGCGAGTCAACCAATGGT 2484
QY 310 LeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLys 329
Db 2485 GCTGATTTGGCCGTATGCTCCGAGAGACTTATTCGTCGGCGGAGAGCACTGCTCTTCGG 2544
QY 330 AlaMetLeuTyrGlnAspIleAlaTyrPheAspGluLysGluAsnSerThrGlyGlyLeu 349
Db 2545 ACGATACTCCGTCAAGACATTCCTTCTTTTGACAGGAAGAGATAGACCGCGCTCTG 2604


```
Qy 350 ThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGly 369
Db 2605 ACCTCTTCCTGTCACCGAGACGAAGCATCTCTCCGGTGTAGGGTGTGACTTAGCC 2664
Qy 370 ValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGly 389
Db 2665 ACGATCTTGATGACCTCCACGACCTTAGAGCGGTATCATATTATGGCCCTCGCATGGG 2724
Qy 390 TrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIle 409
Db 2725 TGGAAATGGCCCTAGTTTGTATCTCGGTGTGCGGTCTCTCGGCATGCGGTTCTTAC 2784
Qy 410 GluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGly 429
Db 2785 CGATTCATATGATAGCCAGTTCATACGCTCCCAAGCTTGCTTATGAGGGATCTGCA 2844
Qy 430 LysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLys 449
Db 2845 AACTTTGCTTCGAGGCTACATCTCTATCCGCACAGTTGCGTCAATTAACCCGGGAAAG 2904
Qy 450 AlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLys 469
Db 2905 GATGCTGGGAGATTACATGCCAGCTTCACGCACAGCAGCAGCATGCTAATCTCT 2964
Qy 470 AlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAla 489
Db 2965 GTCTTGAGGTCACTCCCTGTTATATGCGTCGTCGAGGCACCTGTTTCTCTCGGTTGCG 3024
Qy 490 AlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMet 509
Db 3025 CTCGGGTTTGTGTCGAGGAGACACTCTTGTGTCACCCAGCATGATGACATTTCCCGCTTC 3084
Qy 510 PheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeu 529
Db 3085 TTGTGTTTCTCCGAGATCTCTTTGGTCTCAATCCGGGGCAGCGTCTTTTCTCTTT 3144
Qy 530 AlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLys 549
Db 3145 GCACGACATGGCAGGCGAAGAAATCGCGCCGCGAATTCGACGACGTGTCGACCGA 3204
Qy 550 LysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsn 569
Db 3205 AAGCCACAAATTTGATACTGCTCTGAAGAGGCGGAGAGCTCGAAACGGTGGAAAGGTGAA 3264
Qy 570 LeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArg 589
Db 3265 ATCGAATTTAGAAACGTCCTCTAGATACCCGACCCGCCGACAGCCTGCTCGCGC 3324
Qy 590 GlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCys 609
Db 3325 GGCCTTGACCTGACCGTGAAGCCTGGACAATATGTTGCGCTTGTCCGACCCAGCGGTTGT 3384
Qy 610 GlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnVal 629
Db 3385 GGCAGAGATACCACTTGCATTCCTGAGCGCTTTTACGATGCGATTCGCGGCTCCATC 3444
Qy 630 LeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAla 649
Db 3445 CTGTGTTGATGGAGGACATAGTAAGTAATAATCACTCCTACCGACGCTTCTGTCA 3504
Qy 650 IleValProGlnGluProValPheAsnCysSerIleAlaGluAsnIleAlaTyrGly 669
Db 3505 CTGCTAGCCAGGAGCCGACACTGTACCAGGGCCACCATCAAGGAAACATCTTACTTGGT 3564
Qy 670 AsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIle 689
Db 3565 ATTGTCGAAGATGACGTACCGGAGAAATTTGTGATTAAAGGCTTCGAAGGACGCTAATATC 3624
Qy 690 HisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAla 709
Db 3625 TACACTTCATCATGCTCTCCCGAGGGCTTTAATACAGTTGTGTGAGCAGGAGGAGGC 3684
Qy 710 GlnLeuSerGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysPro 729
```

```
Db 3685 ATGTTGTCTGGCGGCCAAAAGCAACGTTGTGGCCATTGCCGAGCCCTCTCTGGGATCCC 3744
Qy 730 LysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValVal 749
Db 3745 AAAATCCTTCTTCTCGATGAAGCGACGTCAGCCCTCGACTCCGAGTCAGAAAAGGTGTC 3804
Qy 750 GlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeu 769
Db 3805 CAGCGCGCTTGTGATGCGCTGCTCCGAGGCGCAACCAATCGCCGTTCACACCGACTC 3864
Qy 770 SerAlaIleGlnAsnAlaAspLeuIleValLeuHisAsnGlyLysIleLysGluGln 789
Db 3865 AGCACGATTCAAAAGGCGGACGTTATCTATGTTTTCGACCAAGCAAGATCGTCGAAAGC 3924
Qy 790 GlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGln 809
Db 3925 GGAACGACACGCAACTGGTCCAGAAAAGGCGCGGTACTACGAGCTGGTCAACTTCAG 3984
Qy 810 SerVal 811
Db 3985 AGCTTG 3990
```

Search completed: March 31, 2003, 15:23:06
Job time : 266.913 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 31, 2003, 13:52:37 ; Search time 295.562 Seconds
(without alignments)
2337.888 Million cell updates/sec

Title: US-09-873-409-2

Perfect score: 4079

Sequence: 1 MVDENDIPALNVRHYRDHIG.....QELLNRDIYFKLVNAQSVQ 812

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Xgapop 10.0 , Xgapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 574371 seqs, 425486471 residues

Total number of hits satisfying chosen parameters: 1148742

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -DEV=xlh

-MODEL=frame_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09873409/runat_27032003_115424_19290/app_query.fasta_1.7544
-DB=Published Applications NA -QMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09873409 @CGN 1 1 960 @runat_27032003_115424_19290
-NCPUS=6 -ICPU=3 -NO_XLPHY -NO_MMAPP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

- 1: /cgn2_6/prodata/2/pubpna/US07_PUBCOMB.seq:
- 2: /cgn2_6/prodata/2/pubpna/PCT_NEW_PUB.seq:
- 3: /cgn2_6/prodata/2/pubpna/US06_NEW_PUB.seq:
- 4: /cgn2_6/prodata/2/pubpna/US06_PUBCOMB.seq:
- 5: /cgn2_6/prodata/2/pubpna/US07_NEW_PUB.seq:
- 6: /cgn2_6/prodata/2/pubpna/PCTUS_PUBCOMB.seq:
- 7: /cgn2_6/prodata/2/pubpna/US08_NEW_PUB.seq:
- 8: /cgn2_6/prodata/2/pubpna/US08_PUBCOMB.seq:
- 9: /cgn2_6/prodata/2/pubpna/US09_NEW_PUB.seq:
- 10: /cgn2_6/prodata/2/pubpna/US09_PUBCOMB.seq:
- 11: /cgn2_6/prodata/2/pubpna/US10_NEW_PUB.seq:
- 12: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:
- 13: /cgn2_6/prodata/2/pubpna/US60_NEW_PUB.seq:
- 14: /cgn2_6/prodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4079	100.0	2856	10	US-09-873-409-10
2	4079	100.0	3177	10	US-09-873-409-12
3	4079	100.0	3621	10	US-09-873-409-14
4	4079	100.0	3702	10	US-09-873-409-13

5	3456	84.7	2066	10	US-09-873-409-9	Sequence 9, Appli
6	2465.5	60.4	3924	10	US-09-880-107-2299	Sequence 2299, Ap
7	2441	59.8	4369	10	US-09-769-097-1	Sequence 1, Appli
8	2441	59.8	4425	10	US-09-769-097-3	Sequence 3, Appli
9	2438	59.8	4317	9	US-10-044-671-1	Sequence 1, Appli
10	2436	59.7	3860	10	US-09-866-866A-1	Sequence 1, Appli
11	2436	59.7	3860	10	US-09-866-866A-3	Sequence 3, Appli
12	2436	59.7	4643	9	US-10-072-621-2	Sequence 2, Appli
13	2436	59.7	8630	10	US-09-306-417-2	Sequence 2, Appli
14	2436	59.7	8630	10	US-09-306-417-2	Sequence 2, Appli
15	2431.5	59.6	3912	10	US-09-917-800A-1560	Sequence 1560, Ap
16	2425	59.5	4788	10	US-09-866-866A-7	Sequence 7, Appli
17	2404.5	58.9	4189	10	US-09-866-866A-5	Sequence 5, Appli
18	2372	58.2	4254	10	US-09-917-800A-1424	Sequence 1424, Ap
19	2013.5	49.4	5010	10	US-09-917-800A-483	Sequence 483, Appli
20	1694	41.5	4175	10	US-09-749-340-3	Sequence 3, Appli
21	1690.5	41.4	4002	9	US-09-758-828-1	Sequence 1, Appli
22	1690.5	41.4	4002	9	US-09-758-828-3	Sequence 3, Appli
23	1678.5	41.1	3512	10	US-09-749-340-7	Sequence 7, Appli
24	1669.5	40.9	3861	9	US-09-938-842A-2237	Sequence 2237, Ap
25	1617	39.6	2698	10	US-09-749-340-5	Sequence 5, Appli
26	1550.5	38.0	3861	9	US-09-938-842A-263	Sequence 263, App
27	1538.5	37.7	2681	10	US-09-749-340-8	Sequence 8, Appli
28	1452.5	35.6	4653	9	US-10-101-388-2	Sequence 2, Appli
29	1287.5	31.6	8036	9	US-10-101-388-1	Sequence 1, Appli
30	1116.5	27.4	3792	9	US-09-882-694-10	Sequence 10, Appli
31	1032.5	25.3	3999	9	US-09-882-694-9	Sequence 9, Appli
32	866	21.2	7296	10	US-09-070-927A-59	Sequence 59, Appli
33	864	21.2	9047	10	US-09-070-927A-12	Sequence 12, Appli
34	836.5	20.5	1810	10	US-09-749-340-4	Sequence 4, Appli
35	828	20.3	1940	10	US-09-873-409-16	Sequence 16, Appli
36	820.5	20.1	3741	9	US-09-738-626-1251	Sequence 1251, Ap
37	812.5	19.9	2021	10	US-09-873-409-15	Sequence 15, Appli
38	807	19.8	3601	10	US-09-070-927A-255	Sequence 255, App
39	788	19.3	12438	10	US-09-070-927A-173	Sequence 173, App
40	770	18.9	6415	10	US-09-070-927A-275	Sequence 275, Appli
41	769	18.9	640681	10	US-09-790-988-1	Sequence 1, Appli
42	758.5	18.6	2298	9	US-10-156-239-18	Sequence 18, Appli
43	758.5	18.6	2298	10	US-09-795-693-18	Sequence 18, Appli
44	758.5	18.6	3408	9	US-10-156-239-16	Sequence 16, Appli
45	758.5	18.6	3408	10	US-09-795-693-16	Sequence 16, Appli

ALIGNMENTS

RESULT 1
US-09-873-409-10
; Sequence 10, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
; APPLICANT: Frank, Markus
; APPLICANT: Sayegh, Mohamed
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873,409
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
; LENGTH: 2856
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-873-409-10

Alignment Scores:
Pred. No.: 0
Score: 4079.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 10
Length: 2856
Matches: 812
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-09-873-409-2 (1-812) x US-09-873-409-10 (1-2856)

Qy 1 MetValAspGluAsnAspIleAraAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
Db 418 ATGGTGGATGAGATGACATCAGAGCTTTAAATGTGGGCATTAATCGAGACCATATTGGA 477
Qy 21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
Db 478 GTGGTTAGTCAAGAGCCTGTTTGTTCGGGACCAACCATCAGTACAAATATCAAGTATGGA 537
Qy 41 ArgAspAspValThrAspGluGluMetGluuArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
Db 538 CGAGATGATGTCATGCTGAAGAGATGAGAGACAGCAGCAAGGGAAGCAAAATGCGTATGAT 597
Qy 61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyValaGlnMet 80
Db 598 TTTATCATGGAGTTTCCATATAATTTAATACATTTGGTAGGGGAAAAGAGCTCAAAATG 657
Qy 81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaAlaArgAlaLeuValArgAsnProLysIle 100
Db 658 AGTGGAGGCGCAAAACAGAGATCGCAATTCCTGTCCTTTAGTTCGAAACCCCAAGATT 717
Qy 101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120
Db 718 CTGATTTTAGATGAGGCTACGCTCTGCCCTGGATTCAGAAAGCAAGTCAGCTGTTCAAGCT 777
Qy 121 AlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisAtgLeuSerThr 140
Db 778 GCATCTGGAGAAGCGAGCAAAAGGTCGAGTCAATCTGTTGGTAGCACACCAGCTTTCTACT 837
Qy 141 IleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAla 160
Db 838 ATTCGAAGTCAGATTTGATTTGTGACCCCTAAAGGATGGAATGCTGCGGAGAAAGGACA 897
Qy 161 HisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIle 180
Db 898 CATGCTGAACCTAATGGCAAAACGAGGCTATATATTATTCATCTGTGATGTCACAGGATATT 957
Qy 181 LysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSer 200
Db 958 AAAAAAGCTGATGACAGATGGAGTCAATGACATATTCCTACTGAAAGAAAGAACCACTCA 1017
Qy 201 LeuProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSer 220
Db 1018 CTTCTCTGCACCTGTGGAAGAGCATCAAGTCAGACTTTCATTCAGAGGCTGAGGAATCC 1077
Qy 221 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 240
Db 1078 ACCCAATCTAAAGAGATAAGTCTTCTCGAAGTCTCTCTATTAAAAATTTTAAAGTTAAAC 1137
Qy 241 LysProGluTyrProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 260
Db 1138 AAGCCTGAATGGCCTTTTGTGGTTCGGGACATTTGGCTTCTTAAATGGAACTGTT 1197
Qy 261 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLys 280
Db 1198 CATCCAGTATTTCCATCATCTTTGCAAAATTTATACCATGTTTGGAAATAATGATATA 1257
Qy 281 ThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIle 300
Db 1258 ACCACATTAAGACATGATGCAAGAAATTTATTCATGATATTTCGTCATTTTGGGTGTTATT 1317
Qy 301 CysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThr 320
Db 1318 TGCTTTGTGAGTATTTTCATCGAGGATTTATTTACGCGAGAGAGGAGGAAATTTTAAAG 1377
Qy 321 MetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAsp 340
Db 1378 ATGAGATTAAGACACTTGGCCTTCAAGGCCATGTTATATCAGGATATTGCCTGCTTGAT 1437
Qy 341 GluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIle 360
Db 1438 GAAAAGGAAAAACAGCACAGGAGGCTTGACAAATAATTTAGCCATAGATATAGCAAAATT 1497

Qy 361 GlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeu 380
Db 1498 CAAGGAGCAACAGGTTCCAGGATTTGGCGTCTTAACACAAAATGCAACTAATCGGAGCTT 1557
Qy 381 SerValIleIleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAla 400
Db 1558 TCAGTTTATCATTTCCCTTTATATATGATGGAGATGACATTCCTGATTCCTGAGTATTGCT 1617
Qy 401 ProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLys 420
Db 1618 CCAGTACTTCCCGTGACAGGAATGATTGAACCCGACCAATGACTGGATTTGCCAACA 1677
Qy 421 AspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArg 440
Db 1678 GATTAAGCAAGAACTTAAGCATGCTGGAAAGATAGCAACTGAGCTTTGGAGAAATATACGT 1737
Qy 441 ThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGln 460
Db 1738 ACTATAGTGTCAATTAACAGGGAAGAAAGCCTTCGAGCAATGTATGAAGAGATGCTTTCAG 1797
Qy 461 ThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSer 480
Db 1798 ACTCAACACAGAAATACCTCGAAAGAAAGCACAGATTAATGGAAGCTGTTATGCAATTGAGC 1857
Qy 481 HisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGln 500
Db 1858 CATGCCCTTTATATATTTTGGCTATGACAGAGGTTTCGATTTGGAGCTATTAAATTCAA 1917
Qy 501 AlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAla 520
Db 1918 GCTGGACGAATGACCCAGAGGCGATGTTTCATAGTTTTTACTGCAATTCGATATGAGAGCT 1977
Qy 521 MetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAla 540
Db 1978 ATGCCCATCGAAAAAGCGCTCGTTTGGCTCTCGAAATATTCCAAAGCCAAATCGGGGCT 2037
Qy 541 AlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGly 560
Db 2038 GCGCATCTGTTTGGCTTGTGGAAAGAAACCAATATATACAGAGCCGACGTCAGAAAGGG 2097
Qy 561 LysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrPro 580
Db 2098 AAAAAAGCACACACATGTGAAGGAAATTTAGAGTTTTCAGAAAGTCTCTTTCTCTATCCA 2157
Qy 581 CysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThr 600
Db 2158 TGTGCGCCAGATGTTTTCATCCTCCGTTGGCTTATCCCTCAGTATTGAGCGAGGAAGACA 2217
Qy 601 ValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArg 620
Db 2218 GTAGCATTTGTGGGAGCAGCGCTGTGGGAAAAGCACTTCTGTTCAACTTCTGCAGAGA 2277
Qy 621 LeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsn 640
Db 2278 CTTTATGACCCCGTGCAAGGACAAAGTCTGTTGATGGTGGATGCAAAAGAAATTTGAAT 2337
Qy 641 ValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCys 660
Db 2338 GTACAGTGGCTCGTTCCCAAAATAGCAATCGTTCTCCTCAAGAGCTGTGCTCTTCAACTGC 2397
Qy 661 SerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIle 680
Db 2398 AGCATTTGCTGAGAACATCGCTATGTGTGACACAGCCGCTGTGTGCCATTAGATGAGATC 2457
Qy 681 LysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyr 700
Db 2458 AAGAAGCCGCAAAATGACGAAATATCCATTTCTTTATTGAAGGCTCTCCCTGAGAAATAC 2517
Qy 701 AsnThrGlnValGlyLeuLysGlyAlaGluLeuSerGlyGlnLysGlnArgLeuAla 720
Db 2518 AACACAAAGTTGGACTGAAAGGAGCACAGCTTCTTCTGGCGGCGAGAACAAAGAACTAGCT 2577

Qy	721	IleAlaArgAlaLeuLeuGlnIysProIysIleLeuLeuLeuAspGluAlaThrSerAla	740
Db	2578	ATTGCAAGGGCTCTTCTCAAAACCAAAATTTATTGTTGGATGAGCCACATTCAGCC	2637
Qy	741	LeuAspAsnAspSerGluIysValValGlnHisAlaLeuAspLysAlaArgThrGlyArg	760
Db	2638	CTCGATAATGACAGTGAGAGGTGGTTTCAGCATGGCTTGATAAAGCCAGACGGGAAGG	2697
Qy	761	ThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValVal	780
Db	2698	ACATGCTAGTGTGCTACTCACAGGCTCTCTGCAATTCAGAACGCAGATTTTGATAGTGGT	2757
Qy	781	LeuHisAsnGluIysIleIysGluGlnGlyThrHisGlnGluLeuLeuAtqAsnArgAsp	800
Db	2758	CTGCACATGGAAGAATTAAGGNACAGGAGCTCATCAAGAGCTCTCTGAGGAATCGAGAC	2817
Qy	801	IleTyPhylsLeuValAsnAlaGlnSerValGln	812
Db	2818	ATATATTTTAAAGTTAGTGAATGCACAGTCAGTGCAG	2853

RESULT 2

```

US-09-873-409-12
; Sequence 12, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
; APPLICANT: Frank, Markus
; APPLICANT: Savegh, Mohamed
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
; TITLE OF INVENTION: Homologue on Chromosome 7p15-21 and Uses Thereof
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873,409
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 3177
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: No. US20020037522A1e
; LOCATION: (198)..(198)
; OTHER INFORMATION: n at position 198 represents any nucleotide (A, T, C or G)
US-09-873-409-12

```

Alignment Scores:

Pred. NO.:	0	Length:	3177
Alignment Score:	Score:	Matches:	812
	Score:	Conservative:	0
	Percent Similarity:	Mismatches:	0
	Best Local Similarity:	Indels:	0
	Query Match:	Gaps:	0
	DB:		10

US-09-873-409-2 (1-812) x US-09-873-409-12 (1-3177)

Qy	1	MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly	20
Db	739	ATGTTGATGAGAAATGACATCAGAGCTTTAAATGTGCGCATTTATCGAGACCATATTGGA	798
Qy	21	ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly	40
Db	799	GTGGTTAGTCAAGAGCCTGTTTGTTCGGGACCAACATCAGTAACTATCAAGTATGGA	858
Qy	41	ArgAspAspValThrAspGluGluMetGluA:rgAlaAlaArgGluAlaAsnAlaTyrAsp	60
Db	859	CGAGATGATGTGATCGATGAAGAGATGCGAGAGCAGCAGCAAGGGAAGCAAAATGCGCTATGAT	918
Qy	61	PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyValaGlnMet	80
Db	919	TTTATCATGAGATTCTCTAATAATTTTATACATGTTGTTGGGGAAAAAGGAGCTCAATG	978
Qy	81	SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle	100
Db	979	AGTGGGGGCGAGAAACAGAGATCGCAAAATGTGTGTCGCTTAGTTAGTTCGAAACCCCAAGATT	1038

Qy	101	Leu1LeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla	120
Db	1039	CTGATTTTAGATGAGGTACGTCCTCCCTGGATTTCAGAAAGCAAGTCAGCTGTGTTCAAGCT	1098
Qy	121	AlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisAsnLeuSerThr	140
Db	1099	GCACCTGGAGAGCGCAGCAAGGTCGGACTCAATTCGCTGGTAGCACACCGACTTTCCTACT	1158
Qy	141	IleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAla	160
Db	1159	ATTTCGAGTGCAGATTGATTGTGACCTTAAAGGATGGAATGCTGCGCGAGAAAGGACCA	1218
Qy	161	HisAlaGluLeuMetAlaLysArgGlyLeuTyTySerLeuValMetSerGlnAspIle	180
Db	1219	CATGCTGAATTCGCAAAACGAGGCTCTATATTATTTCATTTGATGTCACAGGATATT	1278
Qy	181	LysIleAlaAspGluGlnMetGluSerMetThrTySerThrGluLysGlyThrAsnSer	200
Db	1279	AAAAAGCTGATGAACAGATGAGTCAATATGACATATTTCTCTGAAGAAGAACCAACTCA	1338
Qy	201	LeuProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSer	220
Db	1339	CTTCCTCTGCACCTCTGTGAAGAGATCAAGTCAGACTTCATTGACAGGCTGAGGAATCC	1398
Qy	221	ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn	240
Db	1399	ACCATACTAAAGAGATAAGTCTCTCTGAAGTCTCTCTATTAAAAATTTTAAAGTTAAAC	1458
Qy	241	LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal	260
Db	1459	AAGCCTGAATGGCCTTTTCTGTCTCTGGGGACATTTGGCTCTCTGTTCTAAATTTGGAACCTGTT	1518
Qy	261	HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLys	280
Db	1519	CATCCAGTATTTTCCCATCATCTTTCGAAAAATTTAACCATGTTTGGAAATTAATGATAAA	1578
Qy	281	ThrThrLeuLysHisAspAlaGluIleTySerMetIlePheValIleLeuGlyValIle	300
Db	1579	ACCACATTAAAGCATGATGCAGAAATTTATTCATGATATTTCGTCATTTTGGGTGTTATT	1638
Qy	301	CysPheValSerTyrrPheMetGlnGlyLeuPheTyrrGlyArgAlaGlyGluIleLeuThr	320
Db	1639	TGCTTTGTCAGTTATTTCATGAGGATATTATTTACGCAGACAGCGGCGAAATTTTAACG	1698
Qy	321	MetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrrGlnAspIleAlaThrPheAsp	340
Db	1699	ATGAGATTTAAGACACTTGGCCTTCAAGCCATGTTATATATCAGGATATTTCCTCGTGTGAT	1758
Qy	341	GluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIle	360
Db	1759	GAAGAAGAAAAACACACAGAGGCTTGACNACAATATTAGCCATAGATATAGCACAAATT	1818
Qy	361	GlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeu	380
Db	1819	CAAGGAGCAACAGTTTCCAGGATTCGCGTCTTAAACAAAAATGCAACTAACATCGGACTT	1878
Qy	381	SerValIleIleSerPheIleTyrrGlyTrpGluMetThrPheLeuIleLeuSerIleAla	400
Db	1879	TCAGTTATATCATTTTCCCTTTATATATGATGGGAGATGACATTTCCCTGATTTCTGAGTATTGCT	1938
Qy	401	ProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLys	420
Db	1939	CCAGTACTTCCCTGTACAGGAATGATTGAACCCGACGACATGACTGGATTTTGCCACAAA	1998
Qy	421	AspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArg	440
Db	1999	GATAAGCAAGAACTTAAGCATGCTGGAAGATAGCAACTCTGAAGCTTTTGAGAAATATACGT	2058
Qy	441	ThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrrGluGluMetLeuGln	460
Db	2059	ACTATAGTGTCTAAACAAGGAAAAAGCCTTCGAGCAAAATGTTATGAAGAGATGCTTCAG	2118


```
QY 461 ThrGlnHisArgAsnThrSerLysAlaGlnIlelleGlySerCysTyrAlaPheSer 480
DB 2119 ACTCAACACAGAAATACCTCGAAGAAAGCACAGATTATTGGAAGCTGTATTGCAATTCAGC 2178
QY 481 HisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGln 500
DB 2179 CATCCCTTTATATATTTGGCTATGCGACAGGGTTTCGATTGAGGCTATTAAATTCAA 2238
QY 501 AlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAla 520
DB 2239 GCTGACGAATGACCCAGAGGGCATGTTCTATAGTTTTTACTGCAATTCGATATGAGACT 2298
QY 521 MetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAla 540
DB 2299 ATGGCCATCGAAAAACGCTGTTTGGCTCTCGAATATTCCAAAGCCAAATCGGGGCT 2358
QY 541 AlaHisIlePheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGly 560
DB 2359 GCGCATCTGTTTGGCTTGTGGAAAGAAACCAATATAGACAGCCGCGAGTCAAGAAAGG 2418
QY 561 LysLysProAspThrCysGluGlyValAsnLeuGluPheArgGluValSerPheThrPro 580
DB 2419 AAAAGCCAGACACATGTAAGGGAATTTAGAGTTTCGAGAAAGTCTCTTTCTCTATCCA 2478
QY 581 CysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThr 600
DB 2479 TGTGCGCCAGATGTTTTCATCTCCGTGGCTTATCCCTCAGTATTGAGCGAGGAAGACA 2538
QY 601 ValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArg 620
DB 2539 GTAGCATTTGTTGGGAGCAGCGCTGTGGGAAAGCACTTCTGTTCACATTCCTCAGAGA 2598
QY 621 LeuTyrAspProValGlnGlyValLeuPheAspGlyValAspAlaLysGluLeuAsn 640
DB 2599 CTTTATGACCCCGTGAAGGACAGTCTGTTTGATGTTGATGGTGGATGCAGAAAGATTCGAT 2658
QY 641 ValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCys 660
DB 2659 GTACAGTGGCTCCGTTCCCAATAGCAATCGTTCTCAAGAGCCTGTGCTTCAACTGC 2718
QY 661 SerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIle 680
DB 2719 AGCATTCGTGAGACATCGCTATGTTGGTACACAGCCGCTGTGGTGCCTATTAGATGAGATC 2778
QY 681 LysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyr 700
DB 2779 AAGAGCCGCAATGACGCAATATCATCTCTTTATTTGAAGGTCTCCCTGAGAAATAC 2838
QY 701 AsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAla 720
DB 2839 AACACAAAGTTGACTCAAGAGGACACAGCTTTCTGGCGGCCAGAAACAAAGACTAGCT 2898
QY 721 IleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAla 740
DB 2899 ATTGCAAGGGCTCTTCTCCAAAAACCCAAAAATTTTATTGTTGGATGAGGCCACTTCAGCC 2958
QY 741 LeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArg 760
DB 2959 CTCGATATGACGTGAGAGGTGGTTTCAGATGCCCTTTGATAAAGCCAGACCGGGAAGG 3018
QY 761 ThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValVal 780
DB 3019 ACATGCCCTAGTGGTCACTCAGAGCTCTCTGCAATTCAGAACGCGAGATTGTGATAGTGT 3078
QY 781 LeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAsp 800
DB 3079 CTGCACCAATGGAAGATAAAGGAACAAGGAACCTCATCAAGAGCTCCTGAGAAATCGAGAC 3138
QY 801 IleTyrPheLysLeuValAsnAlaGlnSerValGln 812
DB 3139 ATATATTTTAAAGTTAGTGAATGACAGTCAGTGCAG 3174
```

RESULT 3

US-09-873-409-14

; Sequence 14, Application US/09873409

; Patent No. US20020037522A1

; GENERAL INFORMATION:

; APPLICANT: Frank, Markus

; APPLICANT: Sayegh, Mohamed

; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein

; FILE REFERENCE: 81994/268611

; CURRENT APPLICATION NUMBER: US/09/873,409

; CURRENT FILING DATE: 2001-06-05

; NUMBER OF SEQ ID NOS: 19

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 14

; LENGTH: 3621

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-873-409-14

Alignment Scores:

Pred. No.:	0	Length:	3621
Score:	4079.00	Matches:	812
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	10	Gaps:	0

US-09-873-409-2 (1-812) x US-09-873-409-14 (1-3621)

```
QY 1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
DB 1183 ATGTGTGATGAGAATGACATCAGACGCTTTAAATGTCGGCATTTATCGAGACCATTATTGGA 1242
QY 21 ValValSerGlnGluProValLeuPheGlyThrIleSerAsnAsnIleLysTyrGly 40
DB 1243 GTGGTGTAGTCAAGAGCCTGTTTGTTCGGGACCACCATCAGTAACCAATATCAAGTATGGA 1302
QY 41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
DB 1303 CGAGATGATGCTGACTGATGAAGAGATGGAGAGCAGCAGGAGCAAAATGCGTATGAT 1362
QY 61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80
DB 1363 TTTTATCATGAGTTCCTAATAAATTTAATACATTTGTTAGGGGAAAAAAGAGCTCAAATG 1422
QY 81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100
DB 1423 AGTGGAGGGCAGAAACAGAGGATCGCAATTTGCTGCGCTTAGTTGAAACCCCAAGATT 1482
QY 101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120
DB 1483 CTGATTTTAGTGAAGCTAGCTCTGCGCTGGATTTCAGAAAGCAAGTCAGCTGTTCAAGCT 1542
QY 121 AlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThr 140
DB 1543 GCACCTGGAGAAGCGCAGCAAAAGGTCGAGCTACAATCGTGGTAGCACACCAGCTTCTACT 1602
QY 141 IleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAla 160
DB 1603 ATTCGAAGTCGATTTGATTGTGACCTTAAGGATGGAATGTCGGCGGAGAAAGGAGCA 1662
QY 161 HisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIle 180
DB 1663 CATGCTGAACATAATGCAAAACCGAGGTCTATATTTACTTGTGATGTCACAGGATATT 1722
QY 181 LysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSer 200
DB 1723 AAAAAAGCTGATGAACAGATGGAGTCAATGACATATTTCTACTGAAAAAGAAAGACCACTCA 1782
QY 201 LeuProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSer 220
DB 1783 CTTCCCTCTGCACCTCTGTGAAGAGCATCAAGTCAGACTTTCATTGACAAAGGCTGAGGAATCC 1842
```


Qy	221	ThrGlnSerIysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn	240
Db	1843	ACCCAATCTAAAGAGATAAGTCTCTCTGAAGTCTCTCTATTAAAAAAATTTTAAAGTTTAAAC	1902
Qy	241	LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal	260
Db	1903	AAGCCTGAATGGCTTTTGTGTCTGGGGACATTTGGCTTCTGTCTTAAATGGNACTGTT	1962
Qy	261	HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLys	280
Db	1963	CATCCAGTATTTTCCATCATCTTTCACAAAAATTATAACCATGTTTGGAAATAATGATAAA	2022
Qy	281	ThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIle	300
Db	2023	ACCACATTTAAAGCATGATCAGAAAATTTATTCCATGATATTTCGTCAATTTGGGTGTTATT	2082
Qy	301	CysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThr	320
Db	2083	TGCTTTGTTCAGTTATTTCATGAGGAGTATTATTTACGGCAGACAGGGGAAAATTTTAAAG	2142
Qy	321	MetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAsp	340
Db	2143	ATCAGATTAAAGACACTTGGCCTTCAAGGCCATGTTATATCAGGATATTGCTCGTGTGAT	2202
Qy	341	GluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIle	360
Db	2203	GAAGAAGAAAACAGCAGCAGAGGCTTGACAACAATATTAGCCATAGATATAGACAAAT	2262
Qy	361	GlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeu	380
Db	2263	CAAAGGAGCAACAGTTCAGAGATGGCGTCTTAAACAAAAATGCAACTTAACATGGGACT	2322
Qy	381	SerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAla	400
Db	2323	TCAGTTATCAATTCCTTTATATATGATGGGAGATGACATTCCTGATTCAGATTGCT	2382
Qy	401	ProValLeuAlaValThrGlyMetIleGluThrAlaLaMetThrGlyPheAlaAsnLys	420
Db	2383	CCAGTACTTTGCCGTGACAGAAATGATTGAAACCGCAGCAATGACTGGATTTGCCACAAA	2442
Qy	421	AspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArg	440
Db	2443	GATAAGCAAGAACTTAAGCATCTGTAAGAGATAGCAACTGAAGCTTTGGAGAAATATACGT	2502
Qy	441	ThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGln	460
Db	2503	ACTATAGTGTCTATACAGGGGAAAAGCCTTCGAGCAAAATGATGAGAGATGCTTCAG	2562
Qy	461	ThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSer	480
Db	2563	ACTCAACACAGAAATACCTCGAAGAAAGCACAGATTTATTGGAAGCTGTTATGCATTCAGC	2622
Qy	481	HisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGln	500
Db	2623	CATGCCCTTATATATTTCCTATGCGACAGGGTTTCGATTTGGAGCCCTATTTTAAATCAA	2682
Qy	501	AlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAla	520
Db	2683	GCTGGAGCAATGACCCAGAGGCGATGTTCTAGTATTTTCTACGAAATTCGATATGAGACT	2742
Qy	521	MetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAla	540
Db	2743	ATGGCCATCGAAAAACGCTCGTTTTTGGCTCTCTGAATATTCCAAAGCCAAATTCGGGGCT	2802
Qy	541	AlaHisLeuPheAlaIleuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGly	560
Db	2803	GCGCATCTGTTTGCCTTGTGGAAAAAGAAACCAATATATAGCAGCGCAGATCAAGAAAGG	2862
Qy	561	LysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrPro	580
Db	2863	AAAAGCCAGACACATGTGAAGGGAATTTAGAGTTTCGAAAGTCTCTTTCTTCTATCCA	2922
Qy	581	CysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThr	600

Db	2923	TGTCGCCAGATGTTTTCATCTCCGTGCTTATCCCTCAGTATTGACGAGGAAGACA	2982
Qy	601	ValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArg	620
Db	2983	GTAGCATTTGTGGGAGCAGCGCTGTGGRAAAGCACTTCTGTTCACATTCCTCGACAGA	3042
Qy	621	LeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsn	640
Db	3043	CTTTATGACCCCTGCACAGACAAGTGCTGTTGTGATGTGGATGCAAAAGAATTGAAT	3102
Qy	641	ValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCys	660
Db	3103	GTACAGTGGCTCCGTTCCCAATAGCAATCGTGTCTCAAGAGCCGTGTGCTCTTCACTGCG	3162
Qy	661	SerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIle	680
Db	3163	AGCATTTGCTGAGAACATCGCCTATGGTGACAACACCGCTGTGTGCCATTAGATGAGATC	3222
Qy	681	LysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyr	700
Db	3223	AAAGAAGCCGCAAAATGCAGCAAAATATCCATTCTTTATTGAAGGTCTCCCTGAGAAATAC	3282
Qy	701	AsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAla	720
Db	3283	AACACACAAGTTGGACTGAAGAGGACACAGCTTTCTCGCGGCCAGAAACAAAGACTAGCT	3342
Qy	721	IleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAla	740
Db	3343	ATTGCAAGGGCTCTTCTCCAAAACCCAAAATTTATTGTTGGATGAGCCACTTCAGCC	3402
Qy	741	LeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArg	760
Db	3403	CTCGATAATGACAGTGAGAAGGTGGTTTCAGCATGCCCTTGATAAAGCCAGACGGGAAGG	3462
Qy	761	ThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValVal	780
Db	3463	ACATGCCCTAGTGTGCTACTCACAGGCTCTCTGCAATTCAGAAGCGCAGATTTCATAGTGTT	3522
Qy	781	LeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAsp	800
Db	3523	CTGCACATGTAAGATTAAGGNACAAAGAACTCATCAAGAGCTCTCTGAGAAATCGAGAC	3582
Qy	801	IleTyrPheLysLeuValAsnAlaGlnSerValGln	812
Db	3583	ATATATTTTAAAGTTAGTGAATGCACAGTCAGTCAGTCAG	3618

RESULT 4

```

US-09-873-409-13
; Sequence 13, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
; APPLICANT: Frank, Markus
; APPLICANT: Sayegh, Mohamed
; TITLE OF INVENTION: A Gene Encoding
; TITLE OF INVENTION: Homologue on C
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/87
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 13
; LENGTH: 3702
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: No. US20020037522A1e
; LOCATION: (723)..(723)
; OTHER INFORMATION: n at position 7
US-09-873-409-13

```

Alignment Scores: 0 Length: 3702

Score: 4079.00 Matches: 812
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-873-409-2 (1-812) x US-09-873-409-13 (1-3702)

QY 1 MetValAspGluAsnAspIleAtrGAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
DB 1264 ATGGTGATGAGAAATGACATCAGAGCTTTAAATGTGCGGCAATATCGAGACCAATATTGGA 1323
QY 21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
DB 1324 GTGGTTAGTCAAGAGCCGTGTTTGTTCGGGACCACCATCAGTAACAATAATCAAGTATGGA 1383
QY 41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
DB 1384 CGAGATGATGTGACTGATGAAGAGATGGAGAGCAGCAAGGGAAGCAAAATGCGTATGAT 1443
QY 61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80
DB 1444 TTTATCATGAGATTTCTTAATAATTTAAATACATTTGGTAGGGGAAAAGAGGCTCAATG 1503
QY 81 SerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100
DB 1504 AGTGGAGGCGAGAAACAGAGGATCGCAATTTGCTGCTGCCCTTAGTTTGGAAACCCCAAGATT 1563
QY 101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120
DB 1564 CTGATTTTAGTGGGCTAGCTGCTGCCCTGGATTTCAGAAAGCAAGTCAGCTGTTCAAGCT 1623
QY 121 AlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThr 140
DB 1624 GCATGGAGAGGGGAGCAAGGTGCGACATCAATCTGTGTAGCACACCACCTTTCTACT 1683
QY 141 IleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAla 160
DB 1684 ATTCGAAGTCGAGATTTGATTGTGACCTTAAAGGATGGAATGCTGGCGGAGAAAGAGCA 1743
QY 161 HisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIle 180
DB 1744 CATGCTGAACCTAATGGCAAAACGAGGTCTATATTATTCACTTGTGATGTCACAGGATATT 1803
QY 181 LysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSer 200
DB 1804 AAAAACTGATGAACAGATGGAGTCAATGACATATTCTACTGAAAGAAAGACCAACTCA 1863
QY 201 LeuProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSer 220
DB 1864 CTTCTCTGCACTCTGTGGAAGAGCATCAAGTCAGACTTCATTGACAAGGCTGAGGAATCC 1923
QY 221 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 240
DB 1924 ACCCAATCTAAGAGATAAGTCTTCTCTGAAAGTCTCTCTATTAAATAATTTAAAGTTAAAC 1983
QY 241 LysProGluTrpProPheValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 260
DB 1984 AAGCCTGAATGGCCCTTTGTGGTTCCTGGGACATTTGCTGTCTGTCTTAAATGGAACTGTT 2043
QY 261 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLys 280
DB 2044 CATCCAGTATTTTCCATCACTTTTGGCAAAAATATATAACCATGTTTGGAAATAATGATAAA 2103
QY 281 ThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIle 300
DB 2104 ACCACATTAAGACATGATGCAGAAATTTATTCCATGATATTTCGTCAATTTTGGGTGTTATT 2163
QY 301 CysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThr 320
DB 2164 TGCCTTGTGAGTTATTTTCATGCGGGATTTATTTACGCGAGAGCGGGGAAATTTTAAACG 2223
QY 321 MetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAsp 340

DB 2224 ATCAGATTAAAGCACCTGTCCTTCAAAGCCATGTTATATCAGGATATTCCTCGGTTGAT 2283
QY 341 GluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIle 360
DB 2284 GAAAGGAAACAGACAGAGGCTTGACAACAATATTAGCCATAGATATAGCACAAATT 2343
QY 361 GlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeu 380
DB 2344 CAAGGAGCAACAGGTTCCAGGATTGGCGTCTTAAACACAAAATGCAACTAACATGGGACTT 2403
QY 381 SerValIleIleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAla 400
DB 2404 TCAGTTATCATTTCTTTATATATGATGGAGATGACATTCCTGATTCTGAGTATTGCT 2463
QY 401 ProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLys 420
DB 2464 CAGTACTTCCGTCGACAGGAATGATTGAACCGCAGCAATGACTGGATTTGGCAACAAA 2523
QY 421 AspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArg 440
DB 2524 GATAAGCAAGAACTTAAGCATGCTGGAAGATAGCAACTGAAGCTTTGGAGAATATACGT 2583
QY 441 ThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGln 460
DB 2584 ACTATAGTGTCTATTAAACAAGGAAAGCCCTTCGAGCAAAATGTATGAAGAGATGCTTCAG 2643
QY 461 ThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSer 480
DB 2644 ACTCAACACAGAAATACCTCGAAGAAAGCACAGATTTATTTGGAAAGCTGTTATGCATTCAGC 2703
QY 481 HisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGln 500
DB 2704 CATGCCCTTTATATATTTTGCCTATGCAGCAGGGTTTCGATTTGGAGCCCTATTATTAATCAA 2763
QY 501 AlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAla 520
DB 2764 GCTGGAGCAATCACCCAGAGGCGCATGTTTCATAGTTTCTACTGCAATTCATATGAGCT 2823
QY 521 MetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAla 540
DB 2824 ATGGCCATCGGAAAAACGCTCGTTTTGGCTCTCTGAATATTTCAAAGCCCAATTCGGGGCT 2883
QY 541 AlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGly 560
DB 2884 GCGCATCTGTTTGCCTTGTGGAAAAAGAAACCAAAATATAGACAGCCGAGTCAAGAAGGG 2943
QY 561 LysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrPro 580
DB 2944 AAAAAGCCAGACACATGTGAAGGGAATTTAGAGTTTTCGAGAAGTCTCTTTCTTCTATCCA 3003
QY 581 CysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThr 600
DB 3004 TGTGCGCCAGATGTTTTCATCTCCGTGGCTTATCCCTCAGTATTGAGCGAGGAAGACA 3063
QY 601 ValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArg 620
DB 3064 GTAGCATTTGTGGGAGCAGCGGCTGTGGGAAAAGCACCTTCGTTCAACTTCCTGCAGAGA 3123
QY 621 LeuTyrAspProValGlnGlnValLeuPheAspGlyValAspAlaLysGluLeuAsn 640
DB 3124 CTTTATGACCCCGTCGCAAGGACAAAGTCTGTTTGTGTTGATGGTGGATGCAAAAGAAATTTGAAT 3183
QY 641 ValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCys 660
DB 3184 GTACAGTGGCTCCGTTCCCAAAATAGCAATCGTTCTCAAGAGCTGTGCTCTTCTCAACTGC 3243
QY 661 SerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIle 680
DB 3244 AGCATTTGCTGAGAACATCGCCTATGTGTGACACAGCCGTGTGGTGCATTTAGATGAGATC 3303
QY 681 LysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyr 700

Db 3304 AAGAAGCCCAAAATGAGCAAAATATCCATCTTTTATTGAAGGTCTCCCTCGAGAAATAC 3363
QY 701 AsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAla 720
Db 3364 AACACAAAGTTCGACTGAAAGGAGGACACAGCTTTCTGGCGGCCAGAAAACAAAGACTAGCT 3423
QY 721 ILAAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAla 740
Db 3424 ATTGCAAGGCTCTCTCTCCAAAACCCAAATTTTATTGTTGGATGAGGCCACTTCAGCC 3483
QY 741 LeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArg 760
Db 3484 CTCGTAATACACAGTGAAGAGGGTTCAGCATGCCCTTGTAAAGCCAGGACGGGAAGG 3543
QY 761 ThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValVal 780
Db 3544 ACATGCTAGTGTCTACTCAGGCTCTCTGCAATTCAGAACCGACAGATTGTAGTGGTT 3603
QY 781 LeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAsp 800
Db 3604 CTGCACAAATGGAAGATAAAGGAACCAAGGAACCTCATCAGAGCTCTCGAATAATCGAGAC 3663
QY 801 IleTyrPheLysLeuValAsnAlaGlnSerValGln 812
Db 3664 ATATATTTTAAAGTTAGTGAATGCACAGTCAGTCAG 3699

RESULT 5
US-09-873-409-9
; Sequence 9, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
; APPLICANT: Frank, Markus
; APPLICANT: Sayegh, Mohamed
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
; TITLE OF INVENTION: Homologue on Chromosome 7p15-21 and Uses Thereof
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873,409
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 2066
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-873-409-9

Alignment Scores:
Pred. No.: 0 Length: 2066
Score: 3456.00 Matches: 687
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 84.73% Indels: 0
DB: 10 Gaps: 0

US-09-873-409-2 (1-812) x US-09-873-409-9 (1-2066)

QY 126 SerLysGlyArgThrThrIleValAlaHisArgLeuSerThrIleArgSerAlaAsp 145
Db 3 AGCAAGGTCGGACTACAAATCGTGATGACACACCGACTTCTACTATTCGAAGTGCAGAT 62
QY 146 LeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMet 165
Db 63 TTGATTGTGACCTAAAGGATGGAATCTCGCGGAGAAAGAGACACATGCTGAACTAATG 122
QY 166 AlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIleLysAlaAspGlu 185
Db 123 GCAAAACGAGGCTATATATTACTTGTGATGTACAGGATATTAATAAAGCTGATGAA 182
QY 186 GlnMetCysSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSer 205
Db 183 CAGATGGAGTCAATAGCATATTTCTAGAAAGAAAGCAACCACTTCCTCTGCACTCT 242
QY 206 ValLysSerIleLysSerAspPheIleAspLysAlaGluSerThrGlnSerLysGlu 225

Db 243 GTCAAGAGCATCAAGTCAGACTTCATTGTGCAAGGCTGAGGAAATCCACCAATCTAAAGAG 302
QY 226 IleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTyrPro 245
Db 303 ATAAAGTCTTCCTGAGTCTCTCTATTAAAATTTTAAAGTTAAACAAGCTGATGGCCT 362
QY 246 PheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSer 265
Db 363 TTTGTGGTTCTGGGACATTTGGCTTCTGTTCTAAATGGAACCTGTTTCATCCAGTATTTTC 422
QY 266 IleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThrLeuLysHis 285
Db 423 ATCATCTTTTGCAAAATTAATACCATGTTTGGAAATAATGATAAAACCAATTAAGCAT 482
QY 286 AspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyr 305
Db 483 GATGAGAAATTTATTCATGATATTCGTCATTTGGGTGTTATTTGCTTTGTCAAGTAT 542
QY 306 PheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHis 325
Db 543 TTCATGCGAGGATTTATTTACGCGAGCAGCGGGAATTTTAACGATGAGATTAAGACAC 602
QY 326 LeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGluLysGluAsnSer 345
Db 603 TTTGGCTTCAAAAGCCATGTTATATCAGGATATTTGCTGCTGTTTGAATGAAAAGGAAACAG 662
QY 346 ThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGly 365
Db 663 ACAGGAGGCTTGACAAATATTTAGCCATAGATATAGCACAAATTCAGGAGACACAGGT 722
QY 366 SerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSer 385
Db 723 TCCAGATTTGGGCTCTTAACACAAATGCACTAACTATGGAATTTTCAGTTATCATTTCC 782
QY 386 PheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaVal 405
Db 783 TTTATATATGATGGAGATGACATTCCTCAATTCCTGATTTCTGATTTGCTCCAGTACTTGC 842
QY 406 ThrGlyMetIleGluThrAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeu 425
Db 843 ACAGGAATGATTGAAACCCGACGCAATGACTGGATTTGCCAACAAAGATAGCAAGAACCT 902
QY 426 LysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeu 445
Db 903 AAGCATGCTGGAAGATAGCAACTGAAGCTTTGGAGAAATATAGTACTATAGTGTCTATTA 962
QY 446 ThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsn 465
Db 963 ACAAGGGAAAAAGCCCTTCGAGCAAAATGATGAAGAGATGCTTCAGACTCAACACAGAAAT 1022
QY 466 ThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyr 485
Db 1023 ACCTCGAAGAAAGCAGAGATTTATGGAAGCTGTTATGCAATTCAGCCATGCGCTTATATAT 1082
QY 486 PheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThr 505
Db 1083 TTTGGCTATGACAGCGGTTTCGATTTGGAGCTATTTAATTCAGCTGAGCAATGACC 1142
QY 506 ProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLys 525
Db 1143 CCAGAGGCGCATGTCATAGTTTTTACTGCAATTCATATGAGAGCTATGGCCATCGGAAAA 1202
QY 526 ThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAla 545
Db 1203 ACCTCGTTTGGCTCTCTGAATATTCAAAAGCCAAATCGGGGGCTCGCATCTGTTTGGC 1262
QY 546 LeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThr 565
Db 1263 TTGTTGGAAGAAACCAATATAGACACCGCAGTCAAGAGGGAAGAAAGCCAGACACA 1322
QY 566 CysGluGlyAsnLeuGluPheArgGluValSerPheTyrProCysArgProAspVal 585

Db 1323 TGTGAAGGGAATTTAGAGTTTCGAGAAAGTCTCTTTCTTATCCATGTCGCCCGAGATGTT 1382
QY 586 PheIleuArgGlyLeuSerIleGluArgGlyLysThrValAlaPheValGly 605
Db 1383 TTATCTCTCGTGGCTTATCCCTCAGTATTGAGCGAGAAAGACAGTAGCATTTGTGGGG 1442
QY 606 SerSerGlyCysGlyLysThrSerValGlnLeuLeuGlnArgLeuTyrAspProVal 625
Db 1443 AGCAGCGCTGTGGGAAAGACCTTCTGTTCACCTTCGCAGAGACTTTATGACCCCGTG 1502
QY 626 GlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuValGlnTyrLeuArg 645
Db 1503 CAAGGCAAGTGTCTTTGTGATGGTGTGATGCAAAAGAAATGGAATGACAGTGGCTCCGT 1562
QY 646 SerGlnIleAlaIleValProGlnGluProValLeuPheAsnCyssSerIleAlaGluAsn 665
Db 1563 TCCCAATAGCATCTGTTCTCAAGACCTGTGCTTCTCACTGCAGCATTTGCTGAGAAC 1622
QY 666 IleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAsn 685
Db 1623 ATCGCCTATGGTGACAAACAGCGGTGTGGTGCATTAGATGAGATCAAAAGACCGCAAT 1682
QY 686 AlaAlaSerIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGly 705
Db 1683 GCAGCAATATCCATCTCTTTTATTGAAGGTCTCCCTGAGAAATACAAACACAAAGTTGGA 1742
QY 706 LeuLysGlyAlaGlnLeuSerGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeu 725
Db 1743 CTGAAAGAGCAGACGCTTCTGGCGGCGCAGAAACAGACTAGCTATTGCAAGGGCTCTT 1802
QY 726 LeuGlnLysProtyIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSer 745
Db 1803 CTCCAAAACCCAAATTTTATTGTTGGATGAGGCCACTTCAGCCCTCGATAATGACAGT 1862
QY 746 GluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValVal 765
Db 1863 GAGAAGTGTGTTACAGTGCCTTGTATTAAGCCAGGACGGGAAGACATGCCTAGTGGTC 1922
QY 766 ThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLys 785
Db 1923 ACTCACAGGCTCTCTGCAATTCAGAACGCAGATTTGATGTTCTGCACATGGAAG 1982
QY 786 IleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeu 805
Db 1983 ATAAAGGAACAAGGAACCTCATCAAGAGCTCTCTGAGAAATCGAGACATATATTTAAGTTA 2042
QY 806 ValAsnAlaGlnSerValGln 812
Db 2043 GTGAATGCACAGTCAGTGCAG 2063

RESULT 6

US-09-880-107-2299
; Sequence 2299, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2299
; LENGTH: 3924
; TYPE: DNA
; ORGANISM: Homo sapiens

; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 M23234
US-09-880-107-2299
Alignment Scores: 5,86e-233 Length: 3924
Pred. No.: 2465.50 Matches: 478
Score: 76.46% Conservative: 149
Best Local Similarity: 58.29% Mismatches: 182
Query Match: 60.44% Indels: 11
DB: 10 Gaps: 3
US-09-873-409-2 (1-812) x US-09-880-107-2299 (1-3924)

QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
Db 1395 ATTGATGGCGCAGATATTAGGAACITTAATGTAAACATCTAGGGGAAATCAITTTGGTGTG 1454
QY 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
Db 1455 GTGAGTCAGAGCGGTGCTGTTTCCAGCACAAATTTGCTGAAATATTTTGTATTGGCCGT 1514
QY 42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
Db 1515 GGAATGTAACTATGATGAGATAAAGAAAGCTGTCAAAGAGGCCAACGCTTAIGAGTTT 1574
QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
Db 1575 ATCATGAATATTACCACAGAAATTTGACACCTCTGTTGGAGAGAGAGGGCCAGCTGAGT 1634
QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
Db 1635 GGTGGGCAGAAAGCAGAGGATCGCCATTGACGCTGCTGTTGCGCAACCCCAAGATCCTT 1694
QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
Db 1695 CTGCTGGATGAGCCACGTCAGCAITTTGGACACAGAAAGTGAAGCTGAGTACAGGAGCT 1754
QY 122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
Db 1755 CTGGATAAGCCAGAGAGCGGACCCACCATTTGTATAGCACACCGACTGTCTACGGTC 1814
QY 142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
Db 1815 CGAATATGCAGATGTCATCGCTGGGTTTGAAGATGGAGTAAATTTGTGGAGCAAGAACCCAC 1874
QY 162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuVal----- 175
Db 1875 AGCGAACTGATGAAGAGGAGGGGTGTACTTCAAACCTGTCAACATGCGAGCATCAGGA 1934
QY 176 -----MetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyr 192
Db 1935 AGCCAGATCCAGTCAGAGAAATTTGAACATAATGATGAAAGGCTGCCACTAGATGGCC 1994
QY 193 SerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLys----Ser 211
Db 1995 CCAATGGCTGGAAATCTCGCCTATTAGGCATTTCTACTCAGAAAAACCTTAAAAATTC 2054
QY 212 AspPheIleAspLysAla---GluGluSerThrGlnSerLysGluIleSerLeuProGlu 230
Db 2055 CAAATGTGTGAGAGAGCGCTTGTATGTGGAACCGATGGAGCTTGAAGCAAAATGCGCACCA 2114
QY 231 ValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTyrProPheValValLeuGly 250
Db 2115 GTGCTCTTCTGAAGGCTCTGAAACTGAATAAACAGATGGCCCTACTTTGTGCTGGGA 2174
QY 251 ThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLys 270
Db 2175 ACAGTATGTGCCATTGCCAATGGGGGCTTCAGCCGCGCATTTTCAGTCATATTCTCAGAG 2234
QY 271 IleIleThrMetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyr 290
Db 2235 ATCATAGCGGATTTTGGACCGCGCGATGATGCGTGAAGCAGCAGAGAGTGCACATATTC 2294


```
QY 291 SerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeu 310
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2295 TCTTTCATTTCTTATTCTGGGAATTAATTTCTTTTACTTCTTCTTCCAGGCTTC 2354
QY 311 PheTyrGlyArgAlaGlyGluIleLeuMetArgLeuArgHisLeuAlaPheLysAla 330
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2355 ACCTTTGGGAAGCTGGCAGATCCTCACAGAAAGACTGGCGTCAATGGCTTTAAAGCA 2414
QY 331 MetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThr 350
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2415 ATGCTAAGCAGCAGCATGAGCTGTTTGATGACCAATAAAACAGTACTGGTGCACTTTCT 2474
QY 351 ThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyVal 370
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2475 ACAAGACTTGCCACAGATGCTGCCAAGTCCAAAGGACACAGAAACCAGGTGGGCTTTA 2534
QY 371 LeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrp 390
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2535 ATTGCACAGAATATAGCTAACCTTGGAACTGGTATTATCATATCATTTATCTACGGTTGG 2594
QY 391 GluMetThrPheLeuIleLeuSerIleAlaProValIleAlaValThrGlyMetIleGlu 410
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2595 CAGTTAACCTATTGCTATTAGCAGTTGTTCCAAATTATTGCTGTGTCAGGAATTGTGA 2654
QY 411 ThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGlnLeuLysHisAlaGlyLys 430
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2655 ATGAAATTGTTGGCTGGAATGCAAAAGAGATAAAAGAACTGGAAGCTGCGAAAG 2714
QY 431 IleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAla 450
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2715 ATTGCAACAGAGCAATAGAAAATATTAGACAGATTGTGTCTTTGACCCAGAAAGAAA 2774
QY 451 PheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAla 470
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2775 TTTGAATCATGATGATTGTGAAAAATTGTATGGACCTTACAGGAATTCTGTGCAGAGGCA 2834
QY 471 GlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAla 490
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2835 CACATCTATGGAATTACTTTTAGTATCTCACAGCAATTTATGTATTCTTCTATGCGGT 2894
QY 491 GlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPhe 510
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2895 TGTTTTCGATTTGCTGCATATCTCATTTGTGAATGGACATATGCGCTTCAGAGATGTTAT 2954
QY 511 IleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAla 530
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2955 CTGGTGTCTTCTGCAATTGATTGTTGGTGCAGTGGCTCTAGGACATGCCAGTTCAATTGCT 3014
QY 531 ProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuGluLysLys 550
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3015 CCAGACTATGCTAAGCTAAGCTGTCTGCAGCCCACTTATTCATGCTGTTTGAAGACAA 3074
QY 551 ProAsnIleAspSerArgSerGlnGlyLysLysPheProAspThrCysGluGlyAsnLeu 570
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3075 CCTCTGATTGACAGCTACAGTGAAGAGGGGCTGAAGCCTGATAAATTTGAAGGAATATA 3134
QY 571 GluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGly 590
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3135 ACATTTAATGAATGTTGTTCAACTATCCACCAGCAAAACGTGCCAGTGTTCAGGGG 3194
QY 591 LeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGly 610
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3195 CTGAGCTGGAGGTGAAGAAAGCCAGACACTAGCCCTGGTGGCAGCAGTGGCTGGG 3254
QY 611 LysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeu 630
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3255 AAGAGCACGGTGGTCCAGCTCCTGGAGCGGTCTTACGACCCCTTGGCGGGACAGTGTCT 3314
QY 631 PheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIle 650
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3315 CTCGATGGTCAAGAAAGCAAACTCAATGTCTCAGTGGCTCAGAGCTCAACTCGGAATC 3374
```

```
QY 651 ValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAsp 670
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3375 GTGTCTCAGGAGCCTATCTTATTTGACTGCGACATTTGCCGAGATATTTCCTTATGGAGAC 3434
QY 671 AsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHis 690
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3435 AACAGCGGGTGTATCATCAGGATGAAATTGTGAGTGCAGCCAAAGCTGCCAACATACAT 3494
QY 691 SerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGln 710
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3495 CCTTTCATCGACAGCTTACCCACAAAATATGAAACAAGAGTGGGAGATAAGGGAGCTCAG 3554
QY 711 LeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLys 730
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3555 CTCTCAGGAGGTCAAAAACAGAGGATTTGCTATTGGCCGAGCCCTCATCAGACAACCTCAA 3614
QY 731 IleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGln 750
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3615 ATCTCTCTGTTGGTGAAGCTACATCAGCTCTGGATCTGGAATGGAAGTGAAGGTGTGTCCAA 3674
QY 751 HisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSer 770
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3675 GAAGCCCTGCACAAAGCCAGAGAGCGCGCACCTGCATTGTGATTGCTCACCGCCTGTCC 3734
QY 771 AlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGly 790
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3735 ACCATCCAGAATGCGACTTAATAGTGTGTTTCAGAAATGGGAGAGTCAAGGAGCATGCG 3794
QY 791 ThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 810
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3795 ACGCATCAGCAGCTGCTGCGACAGAAAGGCATCTATTTTCAATGTGCTAGTGTCCAGGCT 3854

RESULT 7
US-09-769-097-1
; Sequence 1, Application US/09769097
; Patent No. US20020055128A1
; GENERAL INFORMATION:
; APPLICANT: Kimberly Anne Brun
; APPLICANT: Richard James Chenery
; APPLICANT: Harma Ellens
; APPLICANT: John Anthony Feild
; APPLICANT: Lin Yue
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES
; FILE OF INVENTION: ENCODING RAT MDR1A AND SCREENING METHODS THEREOF
; FILE REFERENCE: GP-50009-C2
; CURRENT APPLICATION NUMBER: US/09/769,097
; CURRENT FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 09/208,809
; PRIOR FILING DATE: 1998-12-09
; PRIOR APPLICATION NUMBER: 09/156,800
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: US99/20770
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 4369
; TYPE: DNA
; ORGANISM: RATTUS RATTUS
US-09-769-097-1

Alignment Scores:
Pred. No.: 1,79e-230 Length: 4369
Score: 2441.00 Matches: 470
Percent Similarity: 77.04% Conservative: 164
Best Local Similarity: 57.11% Mismatches: 175
Query Match: 59.84% Indels: 14
DB: 10 Gaps: 5

US-09-873-409-2 (1-812) x US-09-769-097-1 (1-4369)
QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```


Db 1684 ATCCAGCGACAGGACATCAGGACCATCAATGTGAGGTATCTGCGGGAATCATTTGGGGTG 1743
Qy 22 ValSerGlnProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
Db 1744 GTGAGTCAGAAACCCGCGCTGTTTGCACACCAATTCGCCAAACATCTCGCTATGCGCGA 1803
Qy 42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
Db 1804 GAAACCTCACCATGGATGAGATAGAGAAAGCTCTCAAGGAGCCATGCTATGATTC 1863
Qy 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
Db 1864 ATCATGAACTGCCCCCAAAATTTGACACCTGTTGCTGGTGGAGAGGGCGCAGCTAGT 1923
Qy 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
Db 1924 GGGGAGCAGAAACAGAGATCGCCATTTGCCCGGCCCTTGTTCGCACACCCCAAGATCCTT 1983
Qy 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
Db 1984 TTGTTGATGAGGCCACGTGAGCTTGCACAGAAAGCGAGCGGTGTTAGGCGCGCT 2043
Qy 122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
Db 2044 CTGGATAAGGCTAGAGAGCGCGACCATTTGTGTGATGCTCACCCGCTTGTCTACAGTT 2103
Qy 142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
Db 2104 CGCAATGCTGACGTCATTTGCTGTTTGTGTTGTTGCTGATGCTGAGAGCAAGGAATCAT 2163
Qy 162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln----- 178
Db 2164 GATGAGCTCATGAGAGAGAAAGGAATTTACTTCAAACTTGTGATGACTCAGACAGCAGGA 2223
Qy 179 ---AspIleLysLysAlaAspGluGlnMetGluSer-----MetThr 191
Db 2224 AATGAAATTTGAATTAGGAAGAAAGGATTTGTAATCTAAAGCAGGAATGATGATGTGCAC 2283
Qy 192 TyrSerThrGluArgLysThrAsnSerLeu-----ProLeuHisSerValLysSerIle 209
Db 2284 ATGTCCTCAAAGAGATTCGGGATCCAGTCTAATAAGAGAGAGATCAACTCGCAAAAGCATC 2343
Qy 210 LysSerAspPheIleAspLysAlaGluGluSerThr---GlnSerLysGluIleSerLeu 228
Db 2344 CGTGGGCCACATGATCAAGACGGGAATTAGCACCAAGAGGCTCTGGATGACAGCGTA 2403
Qy 229 ProGluValSerLeuLysIleLeuLysLeuAsnLysProGluTyrProPheValVal 248
Db 2404 CCTCCAGCTTCCTTTTGGCGGATCCTGAAAGTTGAATTCAACTGAATGGCCTTATTTTGTG 2463
Qy 249 LeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePhe 268
Db 2464 GTTGGTGATTTTGTGCATTAATAATGAGGCTTGCACACGATCTCTCCATNATATTT 2523
Qy 269 AlaLysIleIleThrMetPheLysAsnAspLysThrThrLeu---LysHisAspAla 287
Db 2524 TCRAAGGTGTAGGGTTTTTACAAAAAATGACACCCCTGAAATCCAGCGGACAGACAGC 2583
Qy 288 GluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMet 307
Db 2584 AACTTGTTCCTTTATTTGTTCTGATCCTGGGATCATCTCTTTTCATTACGTTTTTCCTT 2643
Qy 308 GlnGlyLeuPheTyrGlyArgAlaGlyIleLeuThrMetArgLeuArgHisLeuAla 327
Db 2644 CAGGCTTCACATTTGGCAAGCTGGAGAGATCTCTCAACAGGACATCCGATACATGGTC 2703
Qy 328 PheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGluLysGluAsnSerThrGly 347
Db 2704 TTCAAATCCATGCTGAGACAGGACATAGCTGGTTGATGACCTCCCTAAAAACACACAGGA 2763
Qy 348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg 367
Db 2764 CGCGCTGACACAGGCTTGCCTCAATGACGCTCTCAAGTGAAGGGGCTTACAGGGTCTAGG 2823

Qy 368 IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387
Db 2824 CTTGCTGTTTATTACCAGAAACATAGCAAAATCTTGGCAGGCGATCATCATATCCCTGATC 2883
Qy 388 TyrGlyTyrGluMetThrPheLeuLeuSerIleAlaProValLeuAlaValThrGly 407
Db 2884 TAGCGTGGCAATTCACACTTTTACTCTAGCAATTTTCCCATTCATTCCTATACAGGA 2943
Qy 408 MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHis 427
Db 2944 GTGTTGAAATGAAATGTGTCTGGCAAGCGCTGAAAGATAGAGGAAGCTAGAAAGGT 3003
Qy 428 AlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArg 447
Db 3004 TCTGGGAAGATCGCTACAGAGCAATTTGAAACTTTTCGCACCTGCTCTCTTTGACTCGG 3063
Qy 448 GluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSer 467
Db 3064 GAGCAGAAGTTTGAACATATGATGCCAGAGCTTGCAGATACCATACAGAAATGCTTTG 3123
Qy 468 LysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487
Db 3124 AGAAAGCCGACGCTCTTGGGATCACTTCTCTCCACCCAGGCCATGATGATTTCTCC 3183
Qy 488 TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507
Db 3184 TATGCTGCTTGTTCGGTTCGATGCTACTTGGTGGCCAGCACTCATGACATTTGAA 3243
Qy 508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeu 527
Db 3244 AATGTTCTGTTAGTATTTCTCAGCTATTTGTTGGTGCATGGCAGTGGGCGGCTCAGT 3303
Qy 528 ValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeu 547
Db 3304 TCATTCGCTCTGACTACCGGAAAGCCAAAGTCTCGCATCCCATCATCATCAGGATCAT 3363
Qy 548 GluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGlu 567
Db 3364 GAGAAATCCCTGAGATTGACAGCTACAGCAGCGAGGGCTTGAAGCCTAATATGTTGAA 3423
Qy 568 GlyAsnLeuGluPheArgGluValSerPheTyrProCysArgProAspValPheIle 587
Db 3424 GGAATATGAAATTTAATGAGAGTCATGTTCACTATCCACCCGACCCCAATCCCGAGT 3483
Qy 588 LeuArgGlyLysSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSer 607
Db 3484 CTTCAGGGGCTGAGCCTAGAGGTGAAGAAAGGACAGCGCTGGCCCTCTGTTGGGACAGCT 3543
Qy 608 GlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGly 627
Db 3544 GGCTGCGGGAAGAGTACAGTGGTCCAGCTGCTTGAGCGCTTCTATGACCCCATGGCGGA 3603
Qy 628 GlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnThrPleuArgSerGln 647
Db 3604 ACAGTGTTCCTAGATGGCAAGAAATAAAGCAACTCAATGTCAGTGGCTCCCGGCCAC 3663
Qy 648 IleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAla 667
Db 3664 CTGGCATTTGTGCCAGAGGCCCTCTGTTTGTGATGACGATCCCGGAGACATTTGCC 3723
Qy 668 TyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAla 687
Db 3724 TACGGAGACAACAGCGCTGTCGTCTCTATAAGGAGATCGTGAAGGCGAGCCAAAGAGGCC 3783
Qy 688 AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLys 707
Db 3784 AACATCCACAGTTTCATCGACTCACTCGCTGAGAAATACAAACACAGAGTGGGAGACAAA 3843
Qy 708 GlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuGln 727
Db 3844 GGGACTCAGCTGCGGGCGGCGAGAGCAGCGCATGCCCATCGCGCGGCTCTCGTCAGA 3903


```
QY 728 LysProIlyleLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 747
DB 3904 CAGCCTCACATCTTACTTCTGGATGAAGCAGCATCAGCTCTGGATACGGAGTGAAG 3963
QY 748 ValValGlnHisAlaLeuAspLysAlaAgtThrGlyArgThrCysLeuValValThrHis 767
DB 3964 GTCTTCAGGAAGCGTGGCAAGCAGGAGGAGCGGACCTGCATTGTGATCGGCGAC 4023
QY 768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys 787
DB 4024 CGCCTGTCCACCATCCAGAACGCGAGACTTGATCGTGTGATTCAGAACGCGCAGGTCAG 4083
QY 788 GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807
DB 4084 GAGCAGCGGACCCAGCAGCTGCTGGCCAGAAAGGCATCTATTCTCGATGGTCAGT 4143
QY 808 AlaGlnSer 810
DB 4144 GTCCAGGCT 4152
RESULT 8
US-09-769-097-3
; Sequence 3, Application US/09769097
; Patent No. US20020055128A1
; GENERAL INFORMATION:
; APPLICANT: Kimberly Anne Brun
; APPLICANT: Richard James Chenery
; APPLICANT: Harma Ellens
; APPLICANT: John Anthony Feild
; APPLICANT: Lin Yue
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES
; FILE REFERENCE: GP-50009-C2
; CURRENT APPLICATION NUMBER: US/09/769,097
; CURRENT FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 09/208,809
; PRIOR FILING DATE: 1998-12-09
; PRIOR APPLICATION NUMBER: 09/156,800
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: US99/20770
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 4425
; TYPE: DNA
; ORGANISM: RATTUS RATTUS
US-09-769-097-3
Alignment Scores:
Pred. No.: 1,82e-230 Length: 4425
Score: 2441.00 Matches: 470
Percent Similarity: 77.04% Conservatives: 164
Best Local Similarity: 57.11% Mismatches: 175
Query Match: 59.84% Indels: 14
DB: 10 Gaps: 5
US-09-873-409-2 (1-812) x US-09-769-097-3 (1-4425)
QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
DB 1684 ATCCAGCGGACAGGACATCAGGACCATCAATGTGAGGTATCTCGCGGAAATCAATGGGGTG 1743
QY 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
DB 1744 GTGAGTCAGAACCCCGTGTGTTGGCCACCAATTCGCGGAAACATTCGCTATGGCCGA 1803
QY 42 AspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
DB 1804 GAAACGTCACCATGATGATGATAGAGAAAGCTGTTCAGGAGACCATGCTATGATTTC 1863
QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
DB 1944 GTGGTTGAAATGAAATGTTGTCTGGCAAGCGCTGAAAGATAAGAAAGAACTAGAAAGT 3003
```

```
DB 1864 ATCATGAAACTCCCCACAAAATTTGACACCCTGGTGGTGAGAGAGGGCGGAGCTGAGT 1923
QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
DB 1924 GGGGACACAGAAACAGAGGATGCCATTGGCCGGCCCTGTCGCAACCCCAAGATCTTT 1983
QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
DB 1984 TTGTTGGATGAGGCCACGTCAGCTTGGACACAGAAAGCAAGCCGCTGTTTCAGGCCGT 2043
QY 122 LeuGluLysAlaSerLysGlyArgThrThrIleValAlaHisArgLeuSerThrIle 141
DB 2044 CTGGATAGGCTAGAGAGGCGGACCATTTGTGATAGTACCGCTTGTCTACAGTT 2103
QY 142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
DB 2104 CGCAATGCTGAYGTCATTGCTGCTGTTTGTGTTGTGTTGTCATTGTGGAGCAAGAAATCAT 2163
QY 162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln----- 178
DB 2164 GATGAGCTCATGAGAGAGAAAGGAATTTACTTTCAAACCTTGTCTATGACTCAGACGAG 2223
QY 179 ---AspIleLysLysAlaAspGluGlnMetGluSer-----MetThr 191
DB 2224 AATGAAATGAAATAGGAATGAAGCTTGTGAATCTAAAGAYGGAATGATATGATGTGGAC 2283
QY 192 TyrSerThrGluArgLysThrAsnSerLeu-----ProLeuHisSerValLysSerIle 209
DB 2284 ATGTCTTCAAAGATTCRGATCCAGTCTTAATAAGAGAGAGATCAACTCGCAAAAGCATC 2343
QY 210 LysSerAspPheIleAspLysAlaGluGluSerThr---GlnSerLysGluLysSerLeu 228
DB 2344 CGTGGGCCACATGATCAAGACGCGGAACTTAGCACAAAGAGGCTCTGGATCAGCAGTA 2403
QY 229 ProGluValSerLeuLysIleLeuLysLysLeuAsnLysProGluTrpProPheValVal 248
DB 2404 CTCCAGCTTCTCTTTGGCGGATCTGAAAGTTGAAATCAACTGAAATGGGCTTATTTTTGG 2463
QY 249 LeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePhe 268
DB 2464 GTTGGTGTATTTTGTGCCATATAATGAGGCTTGCAACACGACATTTCTCCATATATT 2523
QY 269 AlaLysIleIleThrMetPheGlyAsnAsnAspLysThrLeu---LysHisAspAla 287
DB 2524 TCAAAGGTTGTAGGGGTTTTTACAAAAAATGACACCCCTGAAATCCAGCGGAGAGACAG 2583
QY 288 GluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMet 307
DB 2584 AACTGTTTCTTTTATTTGTTCTGATCTCTGGGATCATCTTTTCAATACGTTTTTCTTT 2643
QY 308 GlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla 327
DB 2644 CAGGCTTCACATTTGGCAAGCTGGAGAGATCTCTCAACAGGACTCCGATACATGGTTC 2703
QY 328 PheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGly 347
DB 2704 TTCAAATCCATGCTGAGACAGGACATAAGCTGTTTGTATGACCTTAAACCAACACAGCA 2763
QY 348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg 367
DB 2764 GCGCTGACACACCGGCTTCCCAATGACGCTGCTCAAGTGAAGGAGGCTACAGGGTCTAG 2823
QY 368 IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387
DB 2824 CTGCTGTTATTACCAGAACATAGCAAAATCTTGGAGCAGGATCATCATATCCCTGATC 2883
QY 388 TyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGly 407
DB 2884 TAGGCTGCAATTTGACACTTTTACTCTTAGCAATTTGTTCCCATCATCTCTATAGCAG 2943
QY 408 MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHis 427
DB 2944 GTGGTTGAAATGAAATGTTGTCTGGCAAGCGCTGAAAGATAAGAAAGAACTAGAAAGT 3003
```



```
QY 428 AlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArg 447
Db 3004 TCTGGGAAGATCCTACAGAAAGCAATGAAACTTTCGCACGTCTCTTTGACTCGG 3063
QY 448 GluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSer 467
Db 3064 GAGCAGAAAGTTTGAAACTATGTATGCCAGAGCTTGCAGATACCACATACAGAAATGCTTGG 3123
QY 468 LysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487
Db 3124 AAGAAAGCGCACGTCTTGGGATCACCTTCTCCTTACCAGGCCATGATGATTTCTCC 3183
QY 488 TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507
Db 3184 TATGCTCTTGTTCGGTTTGATGCTTACTTGTGGCGCAGCACTCATGACATTTGAA 3243
QY 508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeu 527
Db 3244 AATGCTCTGTAGTATCTCAGCTATTGTCTTGTGGTCCATGGCAGTGGCGCAGGTCAGT 3303
QY 528 ValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeu 547
Db 3304 TCATTCTCTCTGACTACGGCAAGCAAAAGTCTCGGCATCCACATCATCAGGATCAT 3363
QY 548 GluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGlu 567
Db 3364 GAGAAATCTCCTGAGATGACAGTACAGCACGGAGGGCTTGAAGCCTATATATGTTGAA 3423
QY 568 GlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIle 587
Db 3424 GGAATGTGAATTTAATGGAGTCATGTTCAACTATCCACCCGACCAACATCCAGTG 3483
QY 588 LeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSer 607
Db 3484 CTTTCAGGGCTGAGCCTAGAGGTGAAGAAAGGCGCAGACGCTGGCCCTCGTGGCGCAGT 3543
QY 608 GlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGly 627
Db 3544 GCGTCCGGGAAGATGACAGTGTCCAGTCTGCTTGGCGCTTCTATGACCCATCGGCCGA 3603
QY 628 GlnValLeuPheAspGlyValAlaAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGln 647
Db 3604 ACATGTTTCTAGATGCAAGAAATAAAGCAACTCAATGTCAGTGGCTCCGGGCCAC 3663
QY 648 IleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAla 667
Db 3664 CTGGGCATTTGTCTCCAGGACCCATCCTGTTGACTGCAGCATCGCCGAGAACATTGCC 3723
QY 668 TyrGlyAspAsnSerArgValProLeuAspGluIleLysGluAlaAlaAsnAlaAla 687
Db 3724 TACGGAGACAAACAGCCGTGTCGTCTCATAGGAGATCGTGAAGGAGCCAGGAGGCC 3783
QY 688 AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLys 707
Db 3784 AACATCCACAGTTCATCGACTCTACTGCTGAGAAATACACACAGAGTGGGAGACAAA 3843
QY 708 GlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuGln 727
Db 3844 GGGACTCAGCTGTCCGGCGGCGAGAGCAGCGCATCGCCATCGCGCGCGCCCTCGTCAGA 3903
QY 728 LysProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 747
Db 3904 CAGCCTTCATCTTACTTCTTGATGAAGCGACATCAGCTCTGGATACGAGAGTGAAG 3963
QY 748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767
Db 3964 GTCGTCCAGGAAGCGCTGACAAAGCCAGGAGGCCCGCACCCTGCAATTGTATCGGCAC 4023
QY 768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys 787
Db 4024 CGCCTGTCCACCATCCAGAGCGCAGACTTGATCGTGGTGATTCAGAACCGGCAGGTCAAG 4083
```

```
QY 788 GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807
Db 4084 GAGCAGCGCACCCAGCAGCTGCTGGCCCAAGAGGCATCTATTTCTCGATGGTCACT 4143
QY 808 AlaGlnSer 810
Db 4144 GTGCAGGCT 4152
RESULT 9
US-10-044-671-1
; Sequence 1, Application US/10044671
; Patent No. US20020177147A1
; GENERAL INFORMATION:
; APPLICANT: Washington State University Research Foundation
; APPLICANT: Mealey, Katrina
; APPLICANT: Bentjen, Steven
; TITLE OF INVENTION: MOR1 VARIANTS AND METHODS FOR THEIR USE
; FILE REFERENCE: 4630-61733
; CURRENT APPLICATION NUMBER: US/10/044,671
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/261,578
; PRIOR FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: US 60/314,829
; PRIOR FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 4317
; TYPE: DNA
; ORGANISM: Canis familiaris
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (70)..(3912)
; OTHER INFORMATION:
US-10-044-671-1
Alignment Scores:
Pred. No.: 3,47e-230 Length: 4317
Score: 2438.00 Matches: 467
Percent Similarity: 76.24% Conservative: 162
Best Local Similarity: 56.61% Mismatches: 178
Query Match: 59.77% Indels: 18
DB: 9 Gaps: 4
US-09-873-409-2 (1-812) x US-10-044-671-1 (1-4317)
QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
Db 1429 ATTTGATGGACAGACATTTAGGACCAATAATGTAAGGCATCTTCGGGAATATTACTGTGTG 1488
QY 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
Db 1489 GTGAGTCAGAGGCTGTGTTGTTTGGCCACCATAGCTGAAACATTCGCTATGGCCG 1548
QY 42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
Db 1549 GAAATATCTCACCATGGATGAGATTGAGAAAGCTGTTAAGGAAGCAATGCTATGATTTT 1608
QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
Db 1609 ATCATGAAACTACTATAATAATTTGACACTCTGTTGGAGAGAGAGGGGCCGCTGAGT 1668
QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
Db 1669 GTTGGACAGAAACAGAGATCGCATTTGCTCGGGCCCTGTTTCCCAACCCAGATTCTT 1728
QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
Db 1729 CTGCTGATCAGGCAACGTCAGCTCTGACACTGAAAGTCAAGCAGTGGTTTCAGTGGCC 1788
QY 122 LeuGlnLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
Db 1789 CTGGATAAGGCCAGAAAGGCGGCGGACTTACCATTTGTGTAGTCTCATCTGTTGTCTACAGTT 1848
```



```

; APPLICANT: Sorrentino, Brian
; APPLICANT: Schuetz, John
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
; FILE REFERENCE: 1340-1-021CIP2
; CURRENT APPLICATION NUMBER: US/09/866,866A
; CURRENT FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 09/584,586
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: PCT/US99/11825
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: 60/086,988
; PRIOR FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 3860
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-866-866A-1

Alignment Scores:
Pred. No.: 4,52e-230 Length: 3860
Score: 2436.00 Matches: 466
Percent Similarity: 76.79% Conservative: 166
Best Local Similarity: 56.62% Mismatches: 177
Query Match: 59.72% Indels: 14
DB: 10 Gaps: 4

US-09-873-409-2 (1-812) x US-09-866-866A-1 (1-3860)

Qy 2 ValAspGluAenAspIleAArgAlaLeuAenValAArgHisIleGlyVal 21
Db 1357 GTTCATGGACAGATATTAGGACCAATTAATGTAAGTTCTTCCGGAAATCATTTGGTGTG 1416

Qy 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAenAenIleLysTyrGlyArg 41
Db 1417 GTGAGTCAGGAACCTGTATTGTTGCCACCACGATAGTCGTAAGAAATCATTCGCTATGGCGGT 1476

Qy 42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAenAlaTyrAspPhe 61
Db 1477 GAAATGTCCACATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAATGCTATGACTTT 1536

Qy 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
Db 1537 ATCATGAAACTGCCTCATAAATTTGACACCTGTGTGGAGAGAGAGGGCCCGCTTGGT 1596

Qy 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValAArgAsnProLysIleLeu 101
Db 1597 GGTGGCAGAGACGAGATCGCATTTGCGACGTGCGCTGTTGCGAACCCCAAGATCCTC 1656

Qy 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
Db 1657 CTGCTGGATGAGGCCACGTGACCTTGGACACAGAAAGCGAGCGTGTTCAGTGGCT 1716

Qy 122 LeuGlyLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
Db 1717 CTGATAGGCCAGAAAAGTTCGACACCACTTGTAGTGTCTCATCGTTGTCTACAGTT 1776

Qy 142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
Db 1777 CGTAATGTCAGCTCATCGCTGGTTTTCGATGATGAGTTCATTGTGGAGAAAGAAATCAT 1836

Qy 162 AlaGluLeuMetAlaLysArgGlyLeuTyrThrSerLeuValMetSerGln----- 178
Db 1837 GATGAACATCATGAAGAAAGGAGGATTTACTTCAAACTTGTCAAAATGCACAGACGAGGA 1896

Qy 179 -----AspIleLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu 195
Db 1897 AATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAATTTGATCGCTTGGAA 1956

Qy 196 ArgLysThrAenSerLeuProLeuHisSerVal-----LysSerIle 209
Db 1957 ATGCTTCAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAAACTCGTAGGAGTGTG 2016

```

```

Qy 210 Lys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeu 228
Db 2017 CGTGGATCAACGCCCAAGACAGACAAAGCTTAGTACCAAGAGGCTCTGGATGAAGTATA 2076

Qy 229 ProGluValSerLeuLysIleLeuLysLeuAenLysProGluTrpProPheValVal 248
Db 2077 CTTCCAGTTTCTCTTTGGAGGATTATGAAGCTAAATTTAACTGAATGGCTTTTGTGTT 2136

Qy 249 LeuGlyThrLeuAlaSerValLeuAenGlyThrValHisProValPheSerIlePhe 268
Db 2137 GTTGGTGTATTTTGTGCCATTATAAATGGAGCGCTGCAACAGCATTGTCAAATAATTT 2196

Qy 269 AlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeuLysHisAspAla 287
Db 2197 TCAAGATTATAGGGGTTTTTACAGAAATTGATGATCCTCAAAACAAACACAGATAGT 2256

Qy 288 GluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMet 307
Db 2257 AACTTGTGTTTCACTATTGTTTCTAGCCCTTGGAAATATTCTTTTATTATCATTTTCTCT 2316

Qy 308 GlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla 327
Db 2317 CAAGGTTTTCATTTTGGCAAAGCTGGAGAGATCCTCACCAGCGGCTCCGATACATGTT 2376

Qy 328 PheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGluLysGluAsnSerThrGly 347
Db 2377 TTCGATCCATGCTCAGACAGAGATGTGAGTTGGTTGATGACCCCTAAACACACCATGGA 2436

Qy 348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg 367
Db 2437 GCATTGACTACCAAGGCTCGCCAAATGATGCTGCTCAAGTTAAAGGGGCTATAGGTTCCAG 2496

Qy 368 IleGlyValLeuThrGlnAenAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387
Db 2497 CTTCGCTGAATACCAGAAATATAGCAAACTCTTGGGACAGAAATAATTATCTCTTCATC 2556

Qy 388 TyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGly 407
Db 2557 TATGTTGGCAACTAACACTGTTACTTCTTAGCAATTGTACCCATCATTTGCAATAGCAGGA 2616

Qy 408 MetIleGluThrAlaAlaMetThrGlyPheAlaAenLysAspLysGlnLeuLysHis 427
Db 2617 GTTGTGAAATGAAATGTTGCTCGACAAAGCTGAAAGATGAAGAAAGAACTAGAGAGT 2676

Qy 428 AlaGlyLysIleAlaThrGluAlaLeuGluAenIleArgThrIleValSerLeuThrArg 447
Db 2677 GCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCGAAACCGTGTGTTCTTTGACTCAG 2736

Qy 448 GluLysAlaPheGluGlnMetTyrGluMetLeuGlnThrGlnHisArgAsnThrSer 467
Db 2737 GACGAGAGTTTGAACATATGATGCTCAGAGTTTGCAGGTACCATACAGAAACTCTTTG 2796

Qy 468 LysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487
Db 2797 AGGAAAGCACACATCTTTGGAATTTACATTTTCTTCCACCCAGGCAATGATGATATTTTC 2856

Qy 488 TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507
Db 2857 TATGCTGGATGTTTCGGTTTGGAGCCTACTTGGTGGCACATAAACTCATGAGCTTTGAG 2916

Qy 508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeu 527
Db 2917 GATGTTCTGTAGTATTCTTGTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2976

Qy 528 ValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeu 547
Db 2977 TCATTTGCTCTGACTATGCAAAAGCCAAATATATCAGCAGCCACATCATGATCATTT 3036

Qy 548 GluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGlu 567
Db 3037 GAAAAAACCCCTTTGATTGACAGTACAGCACGGAAGGCTTAATGCGCAACACACATTGGAA 3096

```



```
QY 568 GlyAsnLeuGluPheArgGluValSerPhePheTyrPrProCysArgProAspValPheIle 587
Db 3097 GGAATGTCACATTTGGTGAAGTTGTTAACTCAATCCACCGACCGGACATCCCGAGT 3156

QY 588 LeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSer 607
Db 3157 CTTGAGGACTGAGCTTGGAGGTGAAGAGGGCCAGACGCTGGCTCTGGTGGGAGGAGT 3216

QY 608 GlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGly 627
Db 3217 GGCTGTGGGAAGAGCACAGTGGTCCAGCTCTCTGGAGCGGTCTTACGACCCCTTTGGCAGG 3276

QY 628 GlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnThrLeuArgSerGln 647
Db 3277 AAAGTGTGCTTGTGTCGCAAGAAATAAGACGACTGAATGTCAGTGGCTCCGAGCACAC 3336

QY 648 IleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAla 667
Db 3337 CTGGGATCGTGTCCAGAGAGCCATCTCTGTTGACTGCGAGCATGCTGAGAACATTTGCC 3396

QY 668 TyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAla 687
Db 3397 TATGGAGACAACAGCCGGTGGTGTCTACAGGAAGAGATCGTGAGGGCAGCAAGAGGCC 3456

QY 688 AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLys 707
Db 3457 AACATACATGCTTCATCGAGTCACTGCTCTAAATAAATATAGCACATAAGTAGGAGACAAA 3516

QY 708 GlyAlaGlnLeuSerGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGln 727
Db 3517 GGAACCTCAGCTCTCTGTGTGTCGCAAGAACACAGCATGCCATGCTGCTGCTTTGTAGA 3576

QY 728 LysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnSerGluLys 747
Db 3577 CAGCCTCATATTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACAGAAAGTGAAG 3636

QY 748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767
Db 3637 GTTGTCCAAAGAAGCCCTGGCAAAAGCCAGAGAGCGGCACCTGCATTTGTGATTCCTAC 3696

QY 768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisGlnGlyLysIleLys 787
Db 3697 CGCCTGTCCACCATCCAGATGACAGCTTAATAGTGTGTGTTCCAGAAATGGCAGTCAAG 3756

QY 788 GlyGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807
Db 3757 GAGCATGGCAGCATGACGAGTCTGCTGGCAGAGAAAGCATCTATTTTCAATGGTCAGT 3816

QY 808 AlaGlnSer 810
Db 3817 GTCAGGCT 3825

RESULT 11
US-09-866-866A-3
; Sequence 3, Application US/09866866A
; Patent No. US20020102244A1
; GENERAL INFORMATION:
; APPLICANT: Sorrentino, Brian
; APPLICANT: Schuetz, John
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
; FILE REFERENCE: 1340-1-021CIP2
; CURRENT APPLICATION NUMBER: US/09/866, 866A
; CURRENT FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 09/584,586
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: PCT/US99/11825
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: 60/086,988
; PRIOR FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 3860
```

```
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-866-866A-3

Alignment Scores:
Pred. No.: 4,62e-230 Length: 3860
Score: 2436.00 Matches: 466
Percent Similarity: 76.79% Conservative: 166
Best Local Similarity: 56.62% Mismatches: 177
Query Match: 59.72% Indels: 14
DB: 10 Gaps: 4

US-09-873-409-2 (1-812) x US-09-866-866A-3 (1-3860)
QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
Db 1357 GTTGATGGACAGATATTAGGACCATAAATGTAAGGTTCTTACGGGAATCATTTGGTGTG 1416

QY 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
Db 1417 GTGAGTCAGGAACCTGTATTGTTTGGCCACCAGATAGCTGAAACACATTCGCTATGGCCGT 1476

QY 42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
Db 1477 GAAATGTGTCACCATGATGAGATTGAGAAAGCTGTCAAGGAAGCCATGCTATGACTTT 1536

QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
Db 1537 ATCATGAAACTGCTCATAAATTTGACACCTGTTGGAGAGAGGGGCCACAGTTGAGT 1596

QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
Db 1597 GGTGGGCAGAAAGCAGAGGATCGCCATTGTCACGTCGCTGTTCCCAACCCCAAGATCCTC 1656

QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
Db 1657 CTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCGTGGTTCAGGTGGCT 1716

QY 122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
Db 1717 CTGGATAAGCCAGAAAGGTGCGACCAACCATTTGTATAGTCTCATCGTTGTCTACAGTT 1776

QY 142 ArgSerAlaAspIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
Db 1777 CGTAATGCTGACGTCATCGCTGCTGTTTCGATGATGAGTCAATTGTGGAGAAAGAAATCAT 1836

QY 162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln----- 178
Db 1837 GATGAATCATGAAGAGAAAGGACATTTACTTCAAACTTGTTCACAAATGCAGCAGCAGGA 1896

QY 179 -----AspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu 195
Db 1897 AATGAATGTAATGATAAATGACGCTGATGATGATCCAAAGTGAATGATGCTTGGAA 1956

QY 196 ArgLysThrAsnSerLeuProLeuHisSerVal-----LysSerIle 209
Db 1957 ATGCTCTCAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACTCGTAGGAGTGC 2016

QY 210 Lys---SerAspPheIleAspLysAlaGluSerThrGlnSerLysGluIleSerLeu 228
Db 2017 CGTGGATCAAGCCCAAGCAGAGAAAGCTAGTACCAAGAGAGGCTCTGGATGAAGATATA 2076

QY 229 ProGluValSerLeuLeuLysIleLeuLysAsnLysProGluTrpProPheValVal 248
Db 2077 CTTCCAGTTTCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGGCTTATTTGTT 2136

QY 249 LeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePhe 268
Db 2137 GTTGGTGTATTTTGTGCCATTATAAATGGAGGCTGCAACACAGCAATTGCAATAATATT 2196

QY 269 AlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeuLysHisAspAla 287
Db 2197 TCAAGATTATGGGGGTTTTTACAGAAATTGATGATCTCTGAAACCAAAACGACAGAAATAGT 2256
```


QY 288 GluIleTyrSerMetIlePheValIleuGlyValIleCysPheValSerTyrPheMet 307
Db 2257 AACTGTGTTTCTACATATGTTTCTAGCCCTTGGAAATATATTTCTTTATATCATTTTTCCTT 2316
QY 308 GlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla 327
Db 2317 CAAGGTTTCACATTTGGCAAGCTGGAGAGATCTCCACCAAGGGCTCCGATCATGTTT 2376
QY 328 PheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlnAsnSerThrGly 347
Db 2377 TTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTGATGACCCCTAAAAACACCACCTGGA 2436
QY 348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg 367
Db 2437 GCATTGACTACAGGCTCGCCCAATGATGCTCAGTTAAAGGGGCTATAGTTTCCAGG 2496
QY 368 IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387
Db 2497 CTTGCTGTGAATTAACCAAGATATAGCAATCTTGGGACAGGAATAATATATATCTTCTATC 2556
QY 388 TyrGlyTrpGluMetThrPheLeuIleuSerIleAlaProValLeuAlaValThrGly 407
Db 2557 TATGGTTGGCAACTAACACTGTTACTTACCAATTTGTACCCATCTTGAATAGCAGGA 2616
QY 408 MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHis 427
Db 2617 GTTGTGAAATGAAATGTTGCTCGGACAGCACTGAAAGATAAGAAAGAACTAGAAAGT 2676
QY 428 AlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArg 447
Db 2677 GCTGGGAAGATCGCTACTGAAGCAATAGAAACTTCCGAACCGTTGTTCTTTGACTCAG 2736
QY 448 GluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSer 467
Db 2737 GAGCAGAAGATTGAACATATATGTCTCAGAGTTTGCAGGTACCATACAGAAACTCTTTG 2796
QY 468 LysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487
Db 2797 AGGAAAGCACACATCTTTGGAAATATACATTTTCTTCCAGGCAATGATGATATTTTCC 2856
QY 488 TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507
Db 2857 TATGCTGGATGTTCCGGTTTGGAGCCTACTTGTGGGCACATAAATCATGAGCTTTGAG 2916
QY 508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeu 527
Db 2917 GATGTTCTGTAGTATTTTACGCTGTTGCTTTGGTGCCATGGCCGTGGGGCAAGTCAGT 2976
QY 528 ValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeu 547
Db 2977 TCATTTGCTCTGATATGCCAAGCCAAATATATCAGCAGGCCCATCATCATCATCAT 3036
QY 548 GluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGlu 567
Db 3037 GAAAAACCCCTTTGATTGACAGCTACAGCAGGAGGAGGCTTAATGTCGGAACACATTGGA 3096
QY 568 GlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIle 587
Db 3097 GGAAATGTACATTTTGGTGAAGTTGTATTCAACTATCCACCCGACCGGACATCCAGTG 3156
QY 588 LeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSer 607
Db 3157 CTTGAGGACTGAGCCTGGAGGTGAAGAGGGCCAGACGCTGGCTCTGTGGGGCAGCAGT 3216
QY 608 GlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGly 627
Db 3217 GCCTGTGGAAGAGACACAGTGTCCAGCTCTCTGGAGCGTTCTACGACCCCTTGGCAGG 3276
QY 628 GlnValLeuPheAspGlyValAspAlaLysGlnLeuAsnValGlnInTrpLeuArgSerGln 647
Db 3277 AAAGTGCTGCTGTGTCGCAAGAAATAAGGACCTGGAATGTTTCAGTTGCTGCCGAGCAC 3336

QY 648 IleAlaIleValProGlnGlnProValLeuPheAsnCysSerIleAlaGluAsnIleAla 667
Db 3337 CTGGGCATCGTGTCCAGGAGCCCATCTGTTTGAAGTCTGAGCATTTGCTGAGAACATTGCC 3396
QY 668 TyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAla 687
Db 3397 TATGGAGACAACAGCCGGTGTGTCTACAGGAGAGATCCTGTAGGGCAGCAAGGAGGCC 3456
QY 688 AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLys 707
Db 3457 AACATACATGCTTTCATCAGTCACTGCTCTTAATAATATAGCACTAAAGTAGGACACAA 3516
QY 708 GlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuGln 727
Db 3517 GGAACCTCAGCTCTCTGTGGCCAGAAACAACGCAATTGCCATAGCTCGTGCCTTGTAGA 3576
QY 728 LysProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 747
Db 3577 CAGCCTCATATTTTGTCTTGGATGAAGCCACGTCAGCTCTGGATACAGAAAGTGAAG 3636
QY 748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767
Db 3637 GTTGTCAAGAGCCCTGGACAAGCCAGAGAAGCGGACCTGCACTGTGATGTGCTCAC 3696
QY 768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys 787
Db 3697 CGCCTGTCCACCATTCCAGAAATGCAGACTTAATAGTGGTGTTCAGAAATGCGAGTCAAG 3756
QY 788 GluGlnGlyThrHisGlnLeuLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807
Db 3757 GAGCATGGCAGCATCAGCAGCTGTGGCAGAAAGGCAATCTATTTTCAATGGTCAGT 3816
QY 808 AlaGlnSer 810
Db 3817 GTCCAGGCT 3825
RESULT 12
US-10-072-621-2
; Sequence 2, Application US/10072621
; Patent No. US20020169137A1
; GENERAL INFORMATION:
; APPLICANT: Reiner, Peter B.
; APPLICANT: Connop, Bruce P.
; APPLICANT: Pollard, Michelle
; TITLE OF INVENTION: REGULATION OF AMYLOID PRECURSOR PROTEIN EXPRESSION
; TITLE OF INVENTION: BY MODIFICATION OF ABC TRANSPORTER EXPRESSION OR ACTIVITY
; FILE REFERENCE: 100103.402
; CURRENT APPLICATION NUMBER: US/10/072,621
; CURRENT FILING DATE: 2002-02-08
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 4643
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-072-621-2
Alignment Scores:
Pred. No.: 6,08e-230 Length: 4643
Score: 2436.00 Matches: 466
Percent Similarity: 76.79% Conservative: 166
Best Local Similarity: 56.62% Mismatches: 177
Query Match: 59.72% Indels: 14
DB: 9 Gaps: 4
US-09-873-409-2 (1-812) x US-10-072-621-2 (1-4643)
QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
Db 1778 GTTGATGGACAGAGATATTAGACCATTAATGTAAAGTTTCTACGGGAATCATTTGGTGTG 1837
QY 22 ValSerGlnGluProValLeuPheGlyThrIleSerAsnAsnIleLysTyrGlyArg 41

Db	1338	GTGAGTCAGGAACCTGTATGTTTGGCCACCAGTAGCTGAAACATTTCGCTATGGCCGT	1897
Qy	42	AspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTy	AspPhe 61
Db	1898	GAAAATGTCACCATGATGAGATTGAGAAGCTGTC AAGGAAGCAATGCCTATGACTTT	1957
Qy	62	IleMetGluPheProLeuLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer	81
Db	1958	ATCATGAAACTGCCTCATAAATTGACACCCCTGGTTGGAGAGAGAGGGGCCAGTTGAGT	2017
Qy	82	GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProValIleLeu	101
Db	2018	GGTGGCGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTGCGAACCCCAAGATCCCT	2077
Qy	102	IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla	121
Db	2078	CTGCTGGATGAGGCCACGTCAGCTTGACACAGAAAGCGAAGCAGTGTTCAGTGGCT	2137
Qy	122	LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle	141
Db	2138	CTGGATAAGCCAGAAAAGGTCCGACCATTTGTGTAGCTCATCGTTTGTCTACAGTT	2197
Qy	142	ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis	161
Db	2198	CGTAATGCTCACGTCTATCGCTGGTTTCGATGATGAGTCACTTTGGAGAAAGAAATCAT	2257
Qy	162	AlaGluLeuMetAlaLysArgGlyLeuTyTYrSerLeuValMetSerGln-----	178
Db	2258	GATGNACTCATGAAGAGAAGAGCATTTACTTCAAACCTGTGCACAATCCAGACAGCAGGA	2317
Qy	179	-----AspIleLysLysAlaAspGluGlnMetGluSerMetThrTYrSerThrGlu	195
Db	2318	AATGAAGTTCAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATTGATGCCTCGAA	2377
Qy	196	ArgLysThrAnSerLeuProLeuHisSerVal-----LysSerIle	209
Db	2378	ATGTCCTTCAAATGATTCACAGATCCAGCTCAANTAAAGAAAAAGATCAACTCGTAGAGAGTGC	2437
Qy	210	Lys---SerAspPheIleAspLysAlaGluSerThrGlnSerLysGluIleSerLeu	228
Db	2438	CGTGGATCACAGCCCAGACAGAAAGCTTAGTACCAAGAGGCTCTCGATGAAAGTATA	2497
Qy	229	ProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpPropheValVal	248
Db	2498	CCTCCAGTTTCCTTTGGAGGATTATGAAGCTAAATTTAACTGAATGSCCTTATTTGTT	2557
Qy	249	LeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePhe	268
Db	2558	GTTGGTGCTATTTGTGCCATTATTAATGAGGCGCTCGCAACCCAGCATTTTTGCAATATATTT	2617
Qy	269	AlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeuLysHisAspAla	287
Db	2618	TCAAGATTTATAGGGTTTTTACAGAATTTGATGATCCTGMAACAAACAGACAGAATAGT	2677
Qy	288	GluIleTySerMetIlePheValIleLeuGlyValIleCysPheValSerTYrPheMet	307
Db	2678	AAC TTGTTTTCATTTGTTTCTAGCCCTTGGAAATTTATTTCTTTTATTACATTTTTCCT	2737
Qy	308	GlnGlyLeuPheTyrglyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla	327
Db	2738	CAGGGTTTCATTTTGGCAAGCTGAGAGATCCCTCAACCAAGCGGCTCCGATACATGGTT	2797
Qy	328	PheLysAlaMetLeuTYrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGly	347
Db	2798	TTCCGATCCCATGCTCAGACAGGATGTGAGTTGGTTTGATGACCCCTAAACCAACCACTGGA	2857
Qy	348	GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg	367
Db	2858	GCATTGACTACCGGCTCGCCATGATGCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGG	2917
Qy	368	IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle	387
Db	2918	CTTGCTGTAATTATCCAGAAATATAGCAAAATCTTGGACAGGAATTAATATATCTTCTATC	2977

Qy	388	TyrGlyTrpGluMetThrPheLeuLeuLeuSerIleAlaProValLeuAlaValThrGly	407
Db	2978	TATGGTTGGCAATAACCATCTGTTACTCTTTAGCAATTGTACCCATCATTTGCAATAGCAGGA	3037
Qy	408	MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHis	427
Db	3038	GTTGTTCAAAATGAANAATGTTGTCTGGCAAGCACTGGAAGAATAGAAGAATAAGAGGT	3097
Qy	428	AlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArg	447
Db	3098	GCTGGGAAGATCGCTACTGAAGCAATAGAAAACCTCCGNAACCGTTGTTCTTTGACTCAG	3157
Qy	448	GluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSer	467
Db	3158	GAGCAGAAGTTTGAACATATGATGCTCAGAGTTTGCAGGTTCAGGTACCATACAGAAATCTTTG	3217
Qy	468	LysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla	487
Db	3218	AGGAAGACACACATCTTTGGAATTAACATTTTCCTTCCACCCAGGCAATGATGATTTTTC	3277
Qy	488	TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu	507
Db	3278	TATGCTGGATGTTTCCGGTTTGAGGCTACTTGTGGGCACATAACTCATGAGCTTTGAG	3337
Qy	508	GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeu	527
Db	3338	GATGTTCTGTTAGTATTTTTCAGCTGTTGCTTTGTGGTCATGGCCGTGGGCAAGTCAGT	3397
Qy	528	ValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeu	547
Db	3398	TCATTTGCTCCTGACTATGCCAAGCCAAATATATCAGCAGGCCACATCATCATGATCATT	3457
Qy	548	GluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGlu	567
Db	3458	GAATAAACCCCTTTTGATTGACAGCTACAGCAGGAAGGCTTAATGCCGCAACACATTGGAA	3517
Qy	568	GlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIle	587
Db	3518	GGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCACCCGACCGGACATCCGAGTG	3577
Qy	588	LeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSer	607
Db	3578	CTTCAGGAGCATGAGCTCGAGGTGAAGAAGGCCACAGACGTGGCTCTCGTGGCAGCAGT	3637
Qy	608	GlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGly	627
Db	3638	GGCTGTGGGAAGACACAGTGGTGCAGCTCTCTGGAGCGGTTCTACGACCCCTTGGCAGGG	3697
Qy	628	GlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGln	647
Db	3698	AAAGTGTGCTTGTATGGCAAGAAATAAAGCCACTGAATGTTTCAGTGGCTCCGAGCACAC	3757
Qy	648	IleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAla	667
Db	3758	CTGGGCATCGTGTCCAGGAGCCATCTCTGTTGACTGCAGCATTCGTCTGAGAACATTGCC	3817
Qy	668	TyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAla	687
Db	3818	TATCGAGACAAACACCGCGGTGGTCTCACAGGAAGAGATTGTGAGGCGCAGCAAGAGGCC	3877
Qy	688	AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValcylLeuLys	707
Db	3878	AACATACATGCTTTCATCGAGTCACTGCCCTATAAATATAGCACTTAAGTAGGAGACAA	3937
Qy	708	GlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuGln	727
Db	3938	GGAACTCAGCTCTCTGGTGGCCAGAAACAACGATTGCCATAGCTCGTGCCCTTGTTAGA	3997
Qy	728	LysProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys	747
Db	3998	CAGCCTCATATTTTGTGTTGATGAAGCCAGTCAGCTCGATCGATACAGAAGTGAAGAAG	4057


```
Qy 748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767
Db 4058 GTTGTCCAGAGAGCCCTGGCAAGCCAGAGAGCGCCACCTGCTGATTGTGCTCAC 4117
Qy 768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys 787
Db 4118 CGCGTGTCCACCATCCAGATGCAGACTTAATAGTGTGTTCAGAAATGGCAGAGTCAAG 4177
Qy 788 GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807
Db 4178 GAGCATGGCAGCATCAGAGCTGCTGGCAGAGAGGAGGAGTCTTAATTTTCAATGGTCACT 4237
Qy 808 AlaGlnSer 810
Db 4238 GTCCAGGCT 4246

RESULT 13
US-09-306-417-1
; Sequence 1, Application US/09306417
; Patent No. US20020103144A1
; GENERAL INFORMATION:
; APPLICANT: Heinrich-Pette-Institut
; TITLE OF INVENTION: Retroviral Gene Transfer Vectors
; FILE REFERENCE: P50491
; CURRENT APPLICATION NUMBER: US/09/306,417
; EARLIER FILING DATE: 1999-05-06
; EARLIER APPLICATION NUMBER: DE 198 22 115
; EARLIER FILING DATE: 1998-05-08
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 8630
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: proviral
; OTHER INFORMATION: plasmid DNA
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(160)
; OTHER INFORMATION: plasmid backbone (pUC)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (161)..(677)
; OTHER INFORMATION: 5'-LTR
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (532)..(1219)
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (1220)..(5062)
; OTHER INFORMATION: m4 mdr-1 cDNA
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (5215)..(5774)
; OTHER INFORMATION: 3'-LTR
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (5775)..(8630)
; OTHER INFORMATION: plasmid backbone (pUC)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(8630)
; OTHER INFORMATION: retroviral expression vector SPbeta71m4
; US-09-306-417-1

Alignment Scores:
Pred. No. : 1.52e-229 Length: 8630
Score: 2435.00 Matches: 466
Percent Similarity: 76.79% Conservatave: 166
Best Local Similarity: 56.62% Mismatches: 177
Query Match: 59.72% Indels: 14
DB: 10 Gaps: 4
```

```
US-09-873-409-2 (1-812) x US-09-306-417-1 (1-8630)
Qy 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
Db 2576 GTTGATGGACAGATATAGGACCATAATATGTAAGTTTCTACGGGAATCATTTGGTGTG 2635
Qy 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
Db 2636 GTGAGTCAGAGAACCTGTAATTTGTCACCAAGTGTGAAACCAATTCGCTATGGCGT 2695
Qy 42 AspAspValThrAspGluMetGluArgAlaAaArgGluAlaAsnAlaTyrAspPhe 61
Db 2696 GAAATGTCAACCATGGATGAGATTGAGAAAGCTGTCAAGGAGCAATGCTATGACTTT 2755
Qy 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
Db 2756 ATCATGAACCTGCTCATAAATTTGACCCCTGTTGGAGAGAGGGGCCCATTTGAGT 2815
Qy 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
Db 2816 GGTGGCAGAGCAGAGGATCGCATTTGCAGTGCCTTGTTCGCAACCCCAAGATCCTC 2875
Qy 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
Db 2876 CTGCTGGATGAGGCCACGTGAGCTTTGGACACAGAAAGCGAGCAGTGTTCAGTGGCT 2935
Qy 122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
Db 2936 CTGATAAGGCCAGAAAGTCGAGCACCATTTGTATAGTCTCATGTTTGTCTACAGTT 2995
Qy 142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
Db 2996 CGTAATGCTGACGTCATCGCTGCTTCGATGATGGAGTCAATTTGGAGAAAGGAATCAT 3055
Qy 162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln----- 178
Db 3056 GATGAACCTCATGAAGAGAAAGGCACTTTACTTCAAACTTGTTCACAAATGTCAGACAGCA 3115
Qy 179 -----AspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu 195
Db 3116 AATGAAGTTGAATAGAAAATGCGATGATGAATCCAAAAGTGCATAAATGATGCTTGGAA 3175
Qy 196 ArgLysThrAsnSerLeuProLeuHisSerVal-----LysSerIle 209
Db 3176 ATGTCCTTCAAAATGATTCACAGATCCAGTCTAATAAGAAAAGATCAACTCGTAGGAGTGC 3235
Qy 210 Lys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeu 228
Db 3236 CGTGGATCAACAAGCCCAAGACAGAAAGCTTAGTACCAGAGGCTCTGGATGAAAGTATA 3295
Qy 229 ProGluValSerLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValVal 248
Db 3296 CCTCCAGTTTCTTTGGAGGATTATGAAGCTAAATTTAACTGAATGGCTTATTTGTT 3355
Qy 249 LeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePhe 268
Db 3356 GTTGTGTATTTTGTGCCATTATAATGGAGGCTGCAACACGAGCATTTGCAATAATATT 3415
Qy 269 AlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeuLysHisAspAla 287
Db 3416 TCAAGATTATAGGGGTTTTTACAGAAATTTGATGATCCTCGAAACAAACACAGAGAAATAGT 3475
Qy 288 GluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMet 307
Db 3476 AACTTGTTTTCACTATTGTTTCTAGCCCTTGGAAATTAATTTCTTTTATATACATTTTCTT 3535
Qy 308 GlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla 327
Db 3536 CAAGTTTTCATTTTGGCAAGCTGGAGAGATCTCTACCAAGCGGCTCCGATACATGTT 3595
Qy 328 PheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGly 347
Db 3476 AACTTGTTTTCACTATTGTTTCTAGCCCTTGGAAATTAATTTCTTTTATATACATTTTCTT
```



```
Db 3596 TTCGATCCATGCTCAGACAGGATGTGAGTTGGTTGATGACCCCTAAACACCACTCGGA 3655
Qy 348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg 367
Db 3656 GCATTGACTACCAAGGCTCGCAATGATGCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGG 3715
Qy 368 IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleSerPheIle 387
Db 3716 CTTGCTGTAATACCCAGAGATATAGCAAAATCTTGGGACAGGAATAATATATCTTCATC 3775
Qy 388 TyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGly 407
Db 3776 TATGGTTGGCAACTAAACACTGTTACTCTTAGCAATTTGACCCATCATTCGCAATAGCAGGA 3835
Qy 408 MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHis 427
Db 3836 GTTGTGTAATGAATGTTGCTGACACAGCACTGAAAGATAAGAAAGAACTAGAAGGT 3895
Qy 428 AlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArg 447
Db 3896 GCTGGGAAGATCGCTACTGAAGCAATAAGAAAATCTCCGAACCGTTGTTCTTGACTCAG 3955
Qy 448 GluLysAlaPheGluGlnMetTyrGluLysMetLeuGlnThrGlnHisArgAsnThrSer 467
Db 3956 GAGCAGAAAGTTTGAACATATATGTATGCTCAGAGTTTGCAGGTACCATCAAGAACTCTTG 4015
Qy 468 LysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487
Db 4016 AGAAAGCACACATCTTTGGAAATACATTTCTTCCACCCAGGCAATGATGATTTTCC 4075
Qy 488 TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507
Db 4076 TATGCTGGATGTTTCCGGTTTGGAGCCTACTTGTGTGCACATAAATCATGAGCTTGAG 4135
Qy 508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeu 527
Db 4136 GATGTTCTGTTAGTATTTTTCAGCTGTTGTTCTTGGTCCCATGCGCGGGGCAAGTCAGT 4195
Qy 528 ValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeu 547
Db 4196 TCATTGCTCTGACTATGTCNAAGCCAAATATATCAGCAGCCCATCATCATCATCAT 4255
Qy 548 GluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGlu 567
Db 4256 GAAAAACCCCTTTGATTGACAGCTACAGCAGGAGGCTTAATGTCGCAACACATTTGAA 4315
Qy 568 GlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIle 587
Db 4316 GGAATGTCATATTGTTGTAAGTTGTATTCAACTATCCACCCGACCGGACATCCCACTG 4375
Qy 588 LeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSer 607
Db 4376 CTTCAGGGACTGAGCCTGGAGGTGAAGAGCGCCAGACGCTGCTGTTGGGGCAGCAGT 4435
Qy 608 GlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGly 627
Db 4436 GGCTGTGGGAAGACACAGTGGTCCAGCTCTCTGGAGCGGTTCTACGACCCCTTGGCAGG 4495
Qy 628 GlnValLeuPheAspGlyValAlaAspAlaLysGluLeuAsnValGlnThrLeuArgSerGln 647
Db 4496 AAGTGTGCTGTGATGGCAAAATAAATTAAGCGACTGAATGTTAGTGGCTCCGAGCACAC 4555
Qy 648 IleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAla 667
Db 4556 CTGGGCATCGTGTCCAGAGGCCCATCTCTGTTGACTGACGATGCTGAGAACTTGCC 4615
Qy 668 TyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAla 687
Db 4616 TATGAGACACACACCGGGTGTGTACAGGAAGAGATCGTGGAGGCGACCAAGAGAGGCC 4675
Qy 688 AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLys 707
Db 4676 AACATACATGCTTTCATCGAGTCTGCTCCTTAATAATATAGCACTAAAGTAGGAGACAAA 4735
```

```
Qy 708 GlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGln 727
Db 4736 GGAACCTCAGCTCTCTGTTGGCCAGAAACAACGATTCCTAGCTCGTCCCTTGTAGA 4795
Qy 728 LysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 747
Db 4796 CAGCCTCATATTGCTTTTGGATGAAGCCAGCTCAGCTCTGGATACAGAAAGTGAAG 4855
Qy 748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767
Db 4856 GTTGTCCAGAGACCTTGACAAAGCCAGAGAGCGCCGACCTGTCATTGTGATTGCTCAC 4915
Qy 768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys 787
Db 4916 CGCCTCTCCACCATCCAGAATGCAGACTTAATAGTGGTCTTTCAGAAATGCAGAGTCAAG 4975
Qy 788 GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807
Db 4976 GAGCATGGCCAGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTTCAATGGTCAGT 5035
Qy 808 AlaGlnSer 810
Db 5036 GTCCAGGCT 5044

RESULT 14
US-09-306-417-2
; Sequence 2, Application US/09306417
; Patent No. US20020103144A1
; GENERAL INFORMATION:
; APPLICANT: Heinrich-Pette-Institut
; TITLE OF INVENTION: Retroviral Gene Transfer Vectors
; FILE REFERENCE: P50491
; CURRENT APPLICATION NUMBER: US/09/306,417
; CURRENT FILING DATE: 1999-05-06
; EARLIER APPLICATION NUMBER: DE 198 22 115
; EARLIER FILING DATE: 1998-05-08
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 8630
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: proviral
; OTHER INFORMATION: plasmid DNA
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(8630)
; OTHER INFORMATION: retroviral expression vector Sfbeta91msA1
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(160)
; OTHER INFORMATION: plasmid backbone (pUC)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (161)..(677)
; OTHER INFORMATION: 5'-LTR
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (532)..(1219)
; FEATURE:
; NAME/KEY: mat peptide
; LOCATION: (1220)..(5062)
; OTHER INFORMATION: msA1 mdrl1 cDNA
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (5215)..(5774)
; OTHER INFORMATION: 3'-LTR
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (5775)..(8630)
; OTHER INFORMATION: plasmid backbone (pUC)
```


US-09-306-417-2

Alignment Scores:

Pred. No.: 1.52e-229 Length: 8630
 Score: 2436.00 Matches: 466
 Percent Similarity: 76.79% Conservative: 166
 Best Local Similarity: 56.62% Mismatches: 177
 Query Match: 59.72% Indels: 14
 DB: 10 Gaps: 4

US-09-873-409-2 (1-812) x US-09-306-417-2 (1-8630)

Qy 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
 Db 2576 GTTGTGACAGGATATTAGGACCATAAATGTAAGGTTCTACGGGAATCATTTGGTGTG 2635
 Qy 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
 Db 2636 GTGAGTCAAGAACCTGTATTGTTGCCACCACGATAGCTGAANAACATTCGCTATGGCCGT 2695
 Qy 42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
 Db 2696 GAAATGTCACATCGATGATGAGAAAGCTGTCAAGGAAGCCATGCCATGACTTT 2755
 Qy 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
 Db 2756 ATCATGAACCTGCCTCATAAATTTGACACCTGGTTGGAGAGAGGGGCCAGTTGAGT 2815
 Qy 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
 Db 2816 GGTGGGCAAGACAGAGATCGCCATTCACGTGCCCTGGTTGCCAACCCTCAAGATCCTC 2875
 Qy 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
 Db 2876 CTGCTGATGAGCCACCTGACCTTGACACAGAAAGCCAGCAGTGTTCAGTGGCT 2935
 Qy 122 LeuGluLysAlaSerLysGlyArgThrThrIleValAlaAlaHisArgLeuSerThrIle 141
 Db 2936 CTGGATAAGGCCAGAAAAGGTCGACCAACCATTTGTATGATGCTCATCGTTTGTCTACAGT 2995
 Qy 142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
 Db 2996 CGTAATGCTGACGTCATCGCTGGTTTCGATGATGAGTCAATTTGGAGAAAGGAATCAT 3055
 Qy 162 AlaGluLeuMetAlaLysArgGlyLeuTyrThrSerLeuValMetSerGln----- 178
 Db 3056 GATGAACTCATGAAGAAAGAGGCAATTTACTTCAAACTTGTCACAACTCAGACAGCAGCA 3115
 Qy 179 -----AspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu 195
 Db 3116 AATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAATTTGATGCTTGAA 3175
 Qy 196 ArgLysThrAsnSerLeuProLeuHisSerVal-----LysSerIle 209
 Db 3176 ATGCTCTCAATGATTCAGATCCAGTCTAATAGAAAAGATCAACTCGTAGGAGTGTCT 3235
 Qy 210 Lys---SerAspPheIleAspLysAlaGluSerThrGlnSerLysGluIleSerLeu 228
 Db 3236 CGTGGATCAAGCCACAGACAGAAAGCTTAGTACCAAGAGGCTCTGGATGAAGTATA 3295
 Qy 229 ProGluValSerLeuLysIleLeuLysLeuAsnLysProGluTyrProPheValVal 248
 Db 3296 CTTCCAGTTTCTTTGGAGGATATGAAGCTAAATTTAACTGAATGGCCTTATTTGTT 3355
 Qy 249 LeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePhe 268
 Db 3356 GTTGGTGTATTGTCGTCATTATTAATGGAGGCTGTCAACCAAGCAGCATTTGCAATTAATTT 3415
 Qy 269 AlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeuLysHisAspAla 287
 Db 3416 TCAAGATTATAGGGTTTTTACAGAAATTGATGATCTCTGAAACAAACAGCAGCAATAGT 3475
 Qy 288 GluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMet 307

Db 3476 AACTGTGTTTCACTATTGTTCTAGCCCTTGGAAATATTCTTTTATTACATTTTCCCTT 3535
 Qy 308 GlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla 327
 Db 3536 CAAGGTTTCACATTTGGCAAGCTGGAGATCTCTACCAAGGGCTCCCATATCATGTT 3595
 Qy 328 PheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGluLysGluAsnSerThrGly 347
 Db 3596 TTCGATCCATGCTCAGACAGGATGTGAGTTGTTGATGACCTAAAAACACCCTGGA 3655
 Qy 348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg 367
 Db 3656 GCATTGACTACCAAGGCTCGCAATGATGCTGTCTCAAGTTAAAGGGGTATAGGTTCAGG 3715
 Qy 368 IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387
 Db 3716 CTTGCTGTAATTAACCAAGATATAGCAATCTTTGGGAGGAGAAATATATCTTCTCATC 3775
 Qy 388 TyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGly 407
 Db 3776 TATGTTGGCAACTAACACACTGTTTACTCTTAGCAATTTGTACCCATCATTTGCAATAGCAGGA 3835
 Qy 408 MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHis 427
 Db 3836 GTTGTGAAATGAAATGTTGTCTGGACAGCACTGAAAGATAAGAAAGAACTAGAAAGT 3895
 Qy 428 AlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArg 447
 Db 3896 GCTGGGAAAGATCGTACTGAAGCAATAGAAAATCTCGAAACCGTTGTTCTTTGACTCAG 3955
 Qy 448 GluLysAlaPheGluGlnMetTyrGluLeuMetLeuGlnThrGlnHisArgAsnThrSer 467
 Db 3956 GAGCAGAAGTTTGAACATATGATGTCTCAGAGTTTGCAGGTACCATACAGAACTCTTTG 4015
 Qy 468 LysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487
 Db 4016 AGGAAAGCACACATCTTTGGAAATTTACATTTCTTCCACCCAGGCAATGATGTATTTTCC 4075
 Qy 488 TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507
 Db 4076 TATGCTGGATGTTTCGGTTTGGAGCCTACTTGTGGCACAATAAATCATGAGCTTTGAG 4135
 Qy 508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeu 527
 Db 4136 GATGTTCTGTGTTAGTATTTTTCAGCTGTTGTTGTTGTCATGCGCGTGGGCAAGTCACT 4195
 Qy 528 ValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeu 547
 Db 4196 TCATTTGCTCTGACTATGCAAGCCAAATAATATCAGCAGCCACCATCATCATGATCAT 4255
 Qy 548 GluLysLysProAsnIleAspSerArgSerGlnGluLysLysProAspThrCysGlu 567
 Db 4256 GAAAAACCCCTTTGATGACGTACAGCAGCAGGAGCCCTAATGCCAACAACATTTGAA 4315
 Qy 568 GlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIle 587
 Db 4316 GGAATGTCATTTGTTGAAGTTGTTTCACTATCCACCCGCGGACATCCCATG 4375
 Qy 588 LeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSer 607
 Db 4376 CTTCCAGGACTGAGCTGGAGGTGAAGAGCCAGACGCTGCTGTTGGGCGAGCAGT 4435
 Qy 608 GlyCysGlyLysSerThrSerValGlnLeuGluArgLeuTyrAspProValGlnGly 627
 Db 4436 GGCTGTGGGAGAGCACAGTGGTCCAGCTCTGAGCGGTTCTACGACCCCTTGGCAGGG 4495
 Qy 628 GlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGln 647
 Db 4496 AAAGTGTCTGCTGATGCGCAAGAAATAAAGCAGCTGAATGTTTCAAGTGGCTCCGAGCAC 4555
 Qy 648 IleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAla 667


```
Db 4556 CTGGGCATCTGTGCTCCAGGAGCCCATCTCTGTTCAGCTGCAGCATTTGCTGAGAACATTGCC 4615
Qy 668 TyxGlyAspAsnSerArgValProLeuAspGluIleLysGluAlaAAsnAlaA 687
Db 4616 TATGGACACAACAGCCGGTGTGTACAGGAAGAGATCGTGAGGGCAGCAAGAGGCC 4675
Qy 688 AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLys 707
Db 4676 AACATACATCGCTTCATCGAGTCACTCGCTTAATAATATATACACTAAAGTAGGAGACAAA 4735
Qy 708 GlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuGln 727
Db 4736 GGAACCTCAGCTCTCTGTGTGCCAGAAACAGCATTCCTGCTGCGCTTGTGTAGA 4795
Qy 728 LysProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 747
Db 4796 CAGCCTCATATTTCCTTTTGGATGAAGCCACGTCAGCTCTGGATACAGAAAGTGAAAG 4855
Qy 748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767
Db 4856 GTTGTCCAGAAAGCCCTGGACAAAGCCAGAGAGCGCCACCTGCATTTGTGATTGCTCAC 4915
Qy 768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys 787
Db 4916 CGCCTGTCCACCATCCAGATGCAGACTTAATAGTGTGTTTCAGAAATGGCAGAGTCAAG 4975
Qy 788 GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807
Db 4976 GAGCATGGCAGCGCATCAGCAGCTGCTGGCACAGNAAGGCATCTATTTTCAATGGTCAGT 5035
Qy 808 AlaGlnSer 810
Db 5036 GTCCAGGCT 5044

RESULT 15
US-09-917-800A-1560
; Sequence 1560, Application US/09917800A
; Patent No. US20020119462A1
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Castle, Arthur
; APPLICANT: Elashoff, Michael
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5038-US
; CURRENT APPLICATION NUMBER: US/09/917, 800A
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/222,040
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: US 60/222,880
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: US 60/290,029
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/290,645
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: US 60/292,336
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/295,798
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/297,457
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,884
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,459
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 1740
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1560
; LENGTH: 3912
; TYPE: DNA
; ORGANISM: Rattus norvegicus
```

```
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020119462A1 NM_012690
US-09-917-800A-1560

Alignment Scores:
Pred. No.: 1,31e-229 Length: 3912
Score: 2431.50 Matches: 470
Percent Similarity: 76.03% Conservative: 155
Best Local Similarity: 57.18% Mismatches: 184
Query Match: 59.61% Indels: 13
DB: 10 Gaps: 2

US-09-873-409-2 (1-812) x US-09-917-800A-1560 (1-3912)
Qy 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
Db 1386 ATCGATGGGCAGGACATCCGGAACCTTTAAAGTTCAGGTGTCTAAGGGAATTCATCGGGCGTG 1445
Qy 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
Db 1446 GTGAGTCAAGAGCGGTACTGTTCTCTACCCAGATTGCTGAAATATATCCGCTATGGCCGT 1505
Qy 42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
Db 1506 GGGAAATGTAACTATGATGAGATTAAAGAGCTGTCAAGAGGCTAATGCTATGACTTC 1565
Qy 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
Db 1566 ATCATGAACTGCCACAGAAATTTGACACCTGTGTTGGTGCAGAGGGGCGCAGCTGAGC 1625
Qy 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
Db 1626 GGGGGACAGAAACAGAGGATCGCCATTGCTGTCCTTGGTCCGCAACCCCAAGATCCTC 1685
Qy 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaA 121
Db 1686 CTGTGGACGAGGCCACGTCAGCCTTTGGACAGAAAGCGAAGCTGAGGTGCAGGCCGCT 1745
Qy 122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
Db 1746 CTGGATAAGCCAGAGAGCGCCGACCAACATCTGTATAGCTCACCCGACTGTCAACTGTC 1805
Qy 142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
Db 1806 CGGAATGCAGATGTCATCGCTGGGTTTGGAGATGGCGTCATCGTGGAGCAAGGAAGCAC 1865
Qy 162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln----- 178
Db 1866 AGTGAGCTGATAAAGAAAGGAGGATCTACTTCAGACTTGTGTAACTGCAGACATCAGGA 1925
Qy 179 -----AspIleLysLysAlaAspGluGlnMetGluSerMet 190
Db 1926 AGCCAGATCCTGTGCAGAGNAATTTGAAGTTGAGCTAAGTATGANAAGGCTGTGGAGGT 1985
Qy 191 ThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLys 210
Db 1986 GTGGCCCCAAATGGCTGGAAAGCAGCGCATATTTAGGAATTTCTACGAAGAAAGTCTGAAA 2045
Qy 211 SerAsp-----PheIleAspLysAlaGluSerThrGlnSerLysGluIleSerLeu 228
Db 2046 AGTTACAGGGCGCATCAAAATAGGCTGGATGGAAACCAATGAACCTTGATCAAAAGCTG 2105
Qy 229 ProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTyrProPheValVal 248
Db 2106 CCACCAAGTGCTTTTCTGAAGGCTCTTAAGCTGAATAAACAACAGATGGCCCTACTTTGTG 2165
Qy 249 LeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePhe 268
Db 2166 GTGGGACACTCTGTGCCATTGCGCAATGGCAAGCGGCCCTCCAGCGCGCATTTCTCCATCATCTG 2225
Qy 269 AlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGlu 288
Db 2226 TCAGAGATGATAGTATCTTTTGGCCCTTGGGGATGACACAGTAAGCAAGAAAGTGTAAAC 2285
```



```
Qy 289 IleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGln 308
Db 2896 ATGTTCTCGCTGCTCTTGGGCTAGGACTCCACTCTCTTTACTTTCTCTTTCAG 2345
Qy 309 GlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPhe 328
Db 2346 GGTTCACATTCGGGAAAGCTGGCGAGATCTCCACCAAGAGCTCCGCTCCATGGCTTC 2405
Qy 329 LysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGly 348
Db 2406 AAGCAATGCTTAAGACAGGACATGAGCTGGTTTACGAGTATATAAAGACGACTGGTGCC 2465
Qy 349 LeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIle 368
Db 2466 CTCCTACAAGACTCGCCACAGAGCGTCGCGAGGTCCAAAGAGGCCACAGGAACCAAGTTG 2525
Qy 369 GlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyr 388
Db 2526 GCTTTAATTGACAGAACACACAGCCAACTTGGAAACGGGTATTATTATATCAATTTATTAC 2585
Qy 389 GlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMet 408
Db 2586 GGTGGCAACTGACACTTCTGCTCTTATCAGTTGTTCATTCATTTGCTGTAGCGGGAATT 2645
Qy 409 IleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAla 428
Db 2646 GTTGAAATGAAATGTTGGCTGGCAAGCCCAAGAGAGATATAAAGGAGATGGAAGCTGCT 2705
Qy 429 GlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGlu 448
Db 2706 GGAAGATTGCAACAGAGGCAATAGAAAATATTGCGACTGTTGTATCTCTTGACCCAAAG 2765
Qy 449 LysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLys 468
Db 2766 AGAAAAATTGAGTCAATGTATGTTGAAAAAATTACCGGACCTTACAGGAATTTCAGTGGG 2825
Qy 469 LysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyr 488
Db 2826 AAGGCTCATCTACCGGATCAGTTTATGATCTCACAAGGATTCATGACTCTTTCTTAT 2885
Qy 489 AlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGly 508
Db 2886 GCTGGCTGCTTTTCGATTGTTGTTCTTACCTCATTTGTAATGGACACATCGGCTTCAAGGAT 2945
Qy 509 MetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuVal 528
Db 2946 GTCATCTCGGTGTTCTACGAATCGTGTGGTGGAGTGGCTCTAGGACATGCCAGCTCA 3005
Qy 529 LeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuGlu 548
Db 3006 TTTGCTCCAGACTATGCAAAAGCCAAAGCTGTCTGCAGCATACTATTATTCAGTCTGTTTGA 3065
Qy 549 LysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGly 568
Db 3066 AGACAACCTCTGATTGACAGCTACAGCAGAGAAGAAATGTGGCCGGAATAAGTTTGAAGA 3125
Qy 569 AsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeu 588
Db 3126 AGCGTGACATTTCAATGAAGTTGTGTTCATTAATTTATCCACCCGGGCCAATGTGCCAGTGCTT 3185
Qy 589 ArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGly 608
Db 3186 CAGGGGCTGAGCCTCGAGTGAAGNAGGGGAGACCCCTGGGCCCTGGTGGGAGTAGTGGC 3245
Qy 609 CysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGln 628
Db 3246 TCGGGAAGAGACCGTGGTCCAGCTGCTCGAGCGCTTTCAGACCCCTTCCAGCCCGGGAACA 3305
Qy 629 ValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIle 648
Db 3306 GTGCTCTCTCGATGGTCAAGGAAGAAAGAAACTCAATGTCTCAGTGGCTCCGAGCTCAACTT 3365
```

```
Qy 649 AlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyr 668
Db 3366 GGCATTGTGTCCTCCAGGAGCCCATCTGTTTGACTGCAGCATCGCAAGAACAATCGCCTAC 3425
Qy 669 GlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAsn 688
Db 3426 GGAACAACAACAGCGTGTGCTGCTCAGGATGAGATTGTGAGGGCGGCCAAGGAGGCCAAC 3485
Qy 689 IleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGly 708
Db 3486 ATCCACCCCTTCATTGAGACACTGCCCCAAAAGTATGATAACAGAGTAGGAGCAAGGGG 3545
Qy 709 AlaGlnLeuSerGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLys 728
Db 3546 ACACAGCTCTCTGGAGGCCAGAAACAGAGGATTGCTATCGCCCGAGCCCTCATCAGACAG 3605
Qy 729 ProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspAsnSerGluLysVal 748
Db 3606 CCTCGGCTCTTACTGTCTGAGTGAAGCCACCTCGGCTTTTGACACTGAGAGTGAAAAGGTC 3665
Qy 749 ValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArg 768
Db 3666 GTCCAGGAAGCGCTGGCAAGGCCAGGAAGGCCGACCTGCAATTGTGATCGCGCACCGC 3725
Qy 769 LeuSerAlaIleGlnAsnAlaAspLeuIleValLeuHisAsnGlyLysIleLysGlu 788
Db 3726 CTGTCCACCATCCAGAACGCGAGACTTGTATCGTGTGATCGACAACGCAAGGTCAGGAG 3785
Qy 789 GlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAla 808
Db 3786 CACGGCACCCACGAGCAGCTGTGGCCCAAGAAAGGCATCTATTCTCCATGGTCAACATT 3845
Qy 809 GlnSer 810
Db 3846 CAAAGCT 3851
```

Search completed: March 31, 2003, 16:00:18

Job time : 444.562 secs

GenCore version 5.1.4 p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 31, 2003, 13:52:37 ; Search time 239.871 Seconds
(without alignments)
2337.888 Million cell updates/sec

Title: US-09-873-409-1

Perfect score: 3323
Sequence: 1 MIAEKAHAELMAKRGLYS.....QELLNRNDIYFKLVNAQSVQ 659

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 574371 seqs, 425486471 residues

Total number of hits satisfying chosen parameters: 1148742

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q/cgn2_1/USPFO.spool/US09873409/runat_27032003_115424_19290/app.query.fasta_1.7544
-DB=Published Applications NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09873409@cgn 1 1.960 @runat_27032003_115424_19290
-NCPU=6 -ICPU=3 -NO_XIPXY -NO_MMAP -LARGQUERY -NEG SCORES=0 -WAIT -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3323	100.0	2066	10	US-09-873-409-9
2	3323	100.0	2856	10	US-09-873-409-10
3	3323	100.0	3177	10	US-09-873-409-12
4	3323	100.0	3621	10	US-09-873-409-14

Alignment Scores:	Pred. NO.:	Length:
Score:	0	2066
Percent Similarity:	3323.00	Matches:
Best Local Similarity:	100.00%	Conservative:
Query Match:	100.00%	Mismatches:
DB:	10	Indels:
		Gaps:

ALIGNMENTS

RESULT 1

US-09-873-409-9
; Sequence 9, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
; APPLICANT: Frank, Markus
; APPLICANT: Sayegh, Mohamed
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
; TITLE OF INVENTION: Homologue on Chromosome 7p15-21 and Uses Thereof
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873.409
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 2066
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-873-409-9

	5	3323	100.0	3702	10	US-09-873-409-13	Sequence 13, Appl
	6	1924.5	57.9	3924	10	US-09-880-107-2299	Sequence 2299, Ap
	7	1882.5	56.7	3912	10	US-09-917-800A-1560	Sequence 1560, Ap
	8	1882	56.6	4369	10	US-09-769-097-1	Sequence 1, Appl
	9	1882	56.6	4425	10	US-09-769-097-3	Sequence 3, Appl
	10	1872	56.3	3860	10	US-09-866-866A-1	Sequence 1, Appl
	11	1872	56.3	3860	10	US-09-866-866A-3	Sequence 3, Appl
	12	1872	56.3	3860	9	US-10-072-621-2	Sequence 2, Appl
	13	1872	56.3	3860	10	US-09-306-417-1	Sequence 1, Appl
	14	1872	56.3	8630	10	US-09-306-417-2	Sequence 2, Appl
	15	1871	56.3	4788	10	US-09-866-866A-7	Sequence 7, Appl
	16	1870	56.3	4317	9	US-10-044-671-1	Sequence 1, Appl
	17	1844.5	55.5	4189	10	US-09-866-866A-5	Sequence 5, Appl
	18	1813	54.6	4254	10	US-09-917-800A-1424	Sequence 1424, Ap
	19	1516.5	45.6	5010	10	US-09-917-800A-483	Sequence 483, Ap
	20	1254	37.7	4002	9	US-09-758-828-1	Sequence 1, Appl
	21	1254	37.7	4002	9	US-09-758-828-3	Sequence 3, Appl
	22	1251.5	37.7	3861	9	US-09-938-842A-2237	Sequence 2237, Ap
	23	1235.5	37.2	2681	10	US-09-749-340-8	Sequence 8, Appl
	24	1232	37.1	4175	10	US-09-749-340-3	Sequence 3, Appl
	25	1226	36.9	2698	10	US-09-749-340-5	Sequence 5, Appl
	26	1211.5	36.5	3512	10	US-09-938-842A-263	Sequence 263, Ap
	27	1098.5	33.1	3861	9	US-10-101-388-2	Sequence 2, Appl
	28	1053.5	31.7	4653	9	US-10-101-388-1	Sequence 1, Appl
	29	840	25.3	3792	9	US-09-882-694-10	Sequence 10, Appl
	30	836.5	25.2	1810	10	US-09-749-340-4	Sequence 4, Appl
	31	828	24.9	1940	10	US-09-873-409-16	Sequence 16, Appl
	32	812.5	24.5	2021	10	US-09-873-409-15	Sequence 15, Appl
	33	783	23.6	3999	9	US-09-882-694-9	Sequence 9, Appl
	34	758.5	22.8	2298	9	US-10-156-239-18	Sequence 18, Appl
	35	758.5	22.8	2298	10	US-09-795-693-18	Sequence 18, Appl
	36	758.5	22.8	3408	9	US-10-156-239-16	Sequence 16, Appl
	37	758.5	22.8	3408	10	US-09-795-693-16	Sequence 16, Appl
	38	758.5	22.8	3512	9	US-10-072-621-1	Sequence 1, Appl
	39	758.5	22.8	3512	10	US-09-815-242-6018	Sequence 6018, Ap
	40	710.5	21.4	1749	10	US-09-815-242-9908	Sequence 9908, Ap
	41	708.5	21.3	1749	10	US-09-815-242-4147	Sequence 4147, Ap
	42	693.5	20.9	1812	10	US-10-260-877-25	Sequence 25, Appl
	43	693	20.9	1764	10	US-09-815-242-6863	Sequence 6863, Ap
	44	649	19.5	2247	9	US-09-870-759-40	Sequence 40, Appl

US-09-873-409-1 (1-659) x US-09-873-409-9 (1-2066)

QY 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
DB 87 ATGCTGGCGGAGAAAGGACCATGCTGAACCTAAATGGCAAAACGAGGCTATATTATTCA 146
QY 21 LeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSer 40
DB 147 CTTGTGATCTCACAGATATTAAAGAGCTGATGAACAGTGGAGTCAATGACATATTCT 206
QY 41 ThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPhe 60
DB 207 ACTGAAAGAAAGAACCAACTCACTCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTC 266
QY 61 IleAspLysAlaGluSerThrGlnSerLysGluLysSerLeuProGluValSerLeu 80
DB 267 ATTGACAAAGGCTGAGGAATCCACCCCAATCTAAAGAGATAAGTCTTCCCTGAAGTCTCTCA 326
QY 81 LeuLysIleLeuLysLeuAsnLysProGluTyrProPheValValLeuGlyThrLeuAla 100
DB 327 TTAAAAATTTAAAGTTAAACCAAGCTGATGGCCCTTTGTGGTCTGGGGACATTTGGCT 386
QY 101 SerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThr 120
DB 387 TCTGTTCTAAATGGAACTGTTTCATCCAGTATTTCATCATCTTTGCAAAAAATTATAACC 446
QY 121 MetPheGlyAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIle 140
DB 447 ATGTTTGGAAATAATGATAAAACCAATTAAAGCATGATGCAGAAATTTATTCATGATA 506
QY 141 PheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGly 160
DB 507 TTCGTCACTTTGGTGCTATTGCTTTGTGCTTATTTTCATGACGGGATTTATTTACGGC 566
QY 161 ArgAlaGlyGluLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyr 180
DB 567 AGACGAGGGGAAATTTTAAACGATGAGATTAAAGACACTTGGCCCTTCAAGACCATGTTATAT 626
QY 181 GlnAspIleAlaTyrPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeu 200
DB 627 CAGGATATTGGCTGGTTTGTATGAAAAGGAAACAGACAGAGGCTTTGACAAATATTA 686
QY 201 AlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGln 220
DB 687 GCCATAGATATAGCAACAATTCAGGAGCAACAGGTTCCAGGATTTGGCGCTTTACACAA 746
QY 221 AsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThr 240
DB 747 AATGCAACTTAACATGGGACTTTTCAGTTATCATTTCTTTATATATGATGGAGATGACA 806
QY 241 PheLeuIleIleuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAla 260
DB 807 TTCCCTGATTTCTGAGTATGCTCCAGTACTTGGCGTGACAGGAATGATGAAACCGCAGCA 866
QY 261 MetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThr 280
DB 867 ATGACTGGATTTGCCAACAAAGATAGCAAGAACTTAAACATGCTGTAAGATAGCAACT 926
QY 281 GluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGln 300
DB 927 GAAGCTTTGGAGAAATATAGCTACTATAGTGTCAATTTCAACAAAGGGAAGCCCTTCGAGCAA 986
QY 301 MetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIle 320
DB 987 ATGTATGAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAACACAGATATT 1046
QY 321 GlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArg 340
DB 1047 GGAAGCTGTATGCACTTACGCCATGCTTTATATATTTTGGCTATCCAGCAGGGTTTCCA 1106
QY 341 PheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPhe 360
DB 1107 TTTGGAGCCTATTATTAATCAAGCTGACGAATGACCCCGAGGGCATGTTTCATAGTTTTT 1166

QY 361 ThrAlaIleAlaTyrGlyValaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyr 380
DB 1167 ACTGCAATTGCATATGAGACTATGGCCATCGAAAAACGCTCGTTTGGCTCCGAATAT 1226
QY 381 SerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIle 400
DB 1227 TCCAAGCCAAATCGGGGCTCGCATCTGTGTGCTTGTGGTGGTGGTGGTGGTGGT 1286
QY 401 AspSerArgSerGlnGluLysLysProAspThrCysGluGlyAsnLeuGluPheArg 420
DB 1287 GACAGCCCGAGTCAAGAAAGGGAAGCCAGACACATGTGAAGGGAATTTAGAGTTTCA 1346
QY 421 GluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeu 440
DB 1347 GAAGTCTCTTCTTCTTCTATCCATGTGCCCCAGATGTTTTTCATCTCCGTTGCTTCCCTC 1406
QY 441 SerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThr 460
DB 1407 AGTATTGAGCGAGGAAAGACAGTAGCATTTTGTGGGAGCAGCGGCTGTGGGAAAAAGCACT 1466
QY 461 SerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGly 480
DB 1467 TCTGTCAACTTCTGCAGAGACTTTTATGACCCCGTGAAGGACAAAGTGTGTTTGTATGT 1526
QY 481 ValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGln 500
DB 1527 GTGGATGCAGAAAGATTCGATGTACAGTGGCTCCGTCCCAATAGCAATCGTTCTCTCAA 1586
QY 501 GluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArg 520
DB 1587 GAGCGTGTGCTCTTCAACTGCAGCATTTCTGAGAAACATCGCCTATGTGTGACAAACGCGT 1646
QY 521 ValValProLeuAspGluIleLysGluAlaAlaAsnIleAlaAsnIleHisSerPheIle 540
DB 1647 GTGGTGCCATTAGATGAGATCAAGAGCCGCAATGAGCAAAATATCCATTCCTTTTAT 1706
QY 541 GluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyValaGlnLeuSerGly 560
DB 1707 GAAGTCTCTCCCTGAGAAATACACACACAGTTGGACTGMAAGGAGCAGACGTTTCTGGC 1766
QY 561 GlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeu 580
DB 1767 GCCCAGAAACAAAGACTAGCTATTGCAAGGCTCTCTCCAAAAACCCAAAAATTTTATTG 1826
QY 581 LeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeu 600
DB 1827 TTGGATGAGGCCACTTCAGCCCTCGATATGACAGTGAAGGTGGTTCAGCATGCCCTT 1886
QY 601 AspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGln 620
DB 1887 GATAAGCCAGGACCGGAGGAGCATGCTTAGTGGTCACTCAGAGCTCTCTGCAATTCA 1946
QY 621 AsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGln 640
DB 1947 AACGCAGATTTGATAGTGTCTGCACAATGGAAGATAAGAAAGAAACAGGAACTCATCAA 2006
QY 641 GluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 659
DB 2007 GAGTCTCTGAGAAATCGAGACATATTTTAAAGTTAGTAGTGAATGCACAGTCAGTCGAC 2063

RESULT 2

US-09-873-409-10

; Sequence 10. Application US/09873409

; Patent No. US20020037522A1

; GENERAL INFORMATION:

; APPLICANT: Frank, Markus

; APPLICANT: Sayegh, Mohamed

; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein

; TITLE OF INVENTION: Homologue on Chromosome 7p15-21 and Uses Thereof

; FILE REFERENCE: 81994/268611

; CURRENT APPLICATION NUMBER: US/09/873,409

; CURRENT FILING DATE: 2001-06-05

; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
; LENGTH: 2856
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-873-409-10

Alignment Scores:

Pred. No.: 0 Length: 2856
Score: 3323.00 Matches: 659
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-873-409-1 (1-659) x US-09-873-409-10 (1-2856)

Qy	1	MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer	20
Db	877	ATGCTGGCGGAGAAAGGACACATGCTGAACCTAATGGCAAAACGAGGTCTATATTATTCA	936
Qy	21	LeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSer	40
Db	937	CTTGTCATGTCACAGGATATTAAAAAAGCTGATGAACAGATGGAGTCAATGACATATTCT	996
Qy	41	ThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPhe	60
Db	997	ACTGNAAGAAAGACCACTCACTTCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTC	1056
Qy	61	IleAspLysAlaGluLysThrGlnSerLysGluIleSerLeuProGluValSerLeu	80
Db	1057	ATTGCAAGGCTGAGGAATCCACCACCACTTAAAGAGATAAGTCTTCTCCTGAAGTCTCTCTA	1116
Qy	81	LeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAla	100
Db	1117	TAAATAATTTTAAAGTTAAACAAGCTGAATGGCGCTTTTGTGGTCTCTGGGGACATTGGCT	1176
Qy	101	SerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThr	120
Db	1177	TCTGTCTTAATGGAACTGTTTCATCCAGTATTTTCATCACTTTTGCACCAAAATTTATACC	1236
Qy	121	MetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIle	140
Db	1237	ATGTTTGGAAATAATGATAAAACACATTTAAAGCATGATGCAGAAATTTATTTCCATGATA	1296
Qy	141	PheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGly	160
Db	1297	TTCTGTCATTTTGGGTGTTATTTTGTCTTGTCTAGTTATTTTCATGCAGGGATTTATTTACGGC	1356
Qy	161	ArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyr	180
Db	1357	AGAGCAGGGGAAATTTTACGATGAGATTAAGACACTTGGCCCTTCAAAGCCATGTTATAT	1416
Qy	181	GlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeu	200
Db	1417	CAGGATATTGCTGTTGATGAAGGAAACACACACAGGAGGCTTGACACAAATATTA	1476
Qy	201	AlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGln	220
Db	1477	GCCATAGATATAGCAAAATTCAGGAGCAACACAGGTTCCAGGATTTGGCGCTCTTAAACAA	1536
Qy	221	AsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThr	240
Db	1537	AATGCAACTTAACATGGGACTTTTCAGTTATCTTTCCTTATATATGGATGGGAGATGACA	1596
Qy	241	PheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAla	260
Db	1597	TTCTCGATTCTGAGTATTGCTCCAGTACTTGGCGTGACAGGAATGATTAAGAACCGCAGCA	1656
Qy	261	MetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThr	280
Db	1657	ATGACTGGATTTGCCCAACAAAGATAAGCAAGAACTTAAGCATGCTGGAAGATAGCAACT	1716

Qy	281	GluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGlyLysAlaPheGluGln	300
Db	1717	GAAGCTTTGGAGAATATACGTACTATAGTGTCAATTAAACAGGAAAAAGCCTTCGAGCAA	1776
Qy	301	MetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIle	320
Db	1777	ATGTATGAAGAGATGCTTCAGACTCAACACAGAATACTCGAAGAAAGACAGATTAAT	1836
Qy	321	GlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArg	340
Db	1837	GGAGCTGTTATGATTCAGCCATGCTTTATATATTTTGCCTATGACAGAGGTTTGA	1896
Qy	341	PheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPhe	360
Db	1897	TTTGAGCCTATTAAATTCAGCTGGACGAATGACCCAGAGGGCATGTTTCATAGTTTTT	1956
Qy	361	ThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyr	380
Db	1957	ACTCAATTTGCATATGGAGCTATGCCATCGGAAAAAGCGCTGTTTGGCTCTCGAATAT	2016
Qy	381	SerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIle	400
Db	2017	TCCAAAGCCAAATCGGGGCTGCGCACTGTTTGGCTTGTGGAAAAAGAACCAATATA	2076
Qy	401	AspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArg	420
Db	2077	GACAGCGCAGTCAAGAAAGGAAAAAGCCACACATGTGAAGGGAATTTAGAGTTTCCA	2136
Qy	421	GluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeu	440
Db	2137	GAAGTCTCTTTCTTATCCATGTGCCCATGTTTTCATCTCCCTCGGTGCTTATCCCTC	2196
Qy	441	SerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThr	460
Db	2197	AGTATTGACCGAGGAACACACAGTAGCATTTGTGGGACAGCGCTGTGGGAAAAAGCACT	2256
Qy	461	SerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGly	480
Db	2257	TCTGTTCAACTTCTGCAGAGACTTTATGACCCCGTGCAAGGACAGTGTGTTGATGCT	2316
Qy	481	ValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGln	500
Db	2317	GTGGATGCAAAAGAAATTGAATGTACAGTGGCTCCGTTCCTCAAAATAGCAATCGTTCTCAA	2376
Qy	501	GluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArg	520
Db	2377	GAGCTGTGCTCTTCACTGCAGCATGCTGAGAACATCGCTATGGTGACAAACAGCGT	2436
Qy	521	ValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIle	540
Db	2437	GTGGTGCCATTTAGATGAGATCAAGAAAGCCCAATGCAAGCAATATCCATTTCTTTATT	2496
Qy	541	GluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGly	560
Db	2497	GAAGGTCTCTCCCTGAGAAATACAAACACAAAGTTGGACTGAAAGGAGACAGCTTTCTGC	2556
Qy	561	GlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuGlnLysProLysIleLeuLeu	580
Db	2557	GGCCAGAAACAAAGACTAGCTATTGCAAGGGCTCTTCTCCAAAAACCCCAAAATTTATTG	2616
Qy	581	LeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeu	600
Db	2617	TTGGATGAGGCCACTTTCAGCCCTCGATTAATGACAGTGAGAGGTGTTTCAGCATGCCCTT	2676
Qy	601	AspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGln	620
Db	2677	GATAAAGCCAGGACGGGAGGACATGCTAGTGGTCACTCAAGGCTCTCTGCAATTCAG	2736
Qy	621	AsnAlaAspLeuIleValValLeuHisGlnGlyLysIleLysGluGlnGlyThrHisGln	640
Db	2737	AACGCAGATTTTGATAGTGGTCTGTCACAAATGGAAAGATAAAGGAAACAGGAACTCATCAA	2796


```
Qy 541 GluGlyLeuProGluLysTyAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGly 560
Db 2818 GAAGGTCTCCCTGAGAAATACACACAAAGTTGGACTGAAAGGAGCACAGCTTTCTGGC 2877
Qy 561 GlyGlnLysGlnArgLeuAlaAlaArgAlaLeuLeuGlnLysProLysIleLeuLeu 580
Db 2878 GGCCAGAAACAAAGACTAGCTATTGCAAGGGCTCTTCTCCAAAACCCAAATTTATTG 2937
Qy 581 LeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeu 600
Db 2938 TTGGATGAGGCCACTTCAGCCCTCGATTAATGACAGTGAAGAGTGGTTTCAGCATGCCCTT 2997
Qy 601 AspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGln 620
Db 2998 GATAAAGCCAGGAGCGGAAGGACATGCCCTAGTGGTCACTCACAGGCTCTCTGCAATTCAG 3057
Qy 621 AsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGln 640
Db 3058 AACGCAGATTTGTAGTGGTTCTGCACAAATGGAAGATAAAGAAACAAGGAACATCATCA 3117
Qy 641 GluLeuLeuArgAsnArgAspIleTyPheLysLeuValAsnAlaGlnSerValGln 659
Db 3118 GAGTCTCTGAGAAATCGAGACATATATTTAAGTTAGTGAATGCACAGTCAAGTGCAG 3174
RESULT 4
US-09-873-409-14
; Sequence 14, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
; APPLICANT: Frank, Markus
; APPLICANT: Sayegh, Mohamed
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human p-Glycoprotein
; TITLE OF INVENTION: Homologue on Chromosome 7p15-21 and Uses Thereof
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873,409
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
; LENGTH: 3621
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-873-409-14
Alignment Scores:
Pred. No.: 0 Length: 3621
Score: 3323.00 Matches: 659
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0
US-09-873-409-1 (1-659) x US-09-873-409-14 (1-3621)
Qy 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyPyrSer 20
Db 1642 ATGTGCGGAGAAAGGAGCACATCTGAACCTAATGGCAAAACGAGTCTATATTATTC 1701
Qy 21 LeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyPyrSer 40
Db 1702 CTTGTGTGTCTACAGGATATTAATAAGCTGATGAACAGATGGAGTCAATGACATATTTCT 1761
Qy 41 ThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPhe 60
Db 1762 ACTGAAAGAAAGAACCAACTCTCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTC 1821
Qy 61 IleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeu 80
Db 1822 ATTGACAGGCTGAGGAATCCACCACCAATCAAGAGATPAAGTCTTCTCCTGAAGTCTCTCTA 1881
Qy 81 LeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAla 100
Db 1882 TTAAAAATTTTAAAGTTAAACAAGCTCGAATGGCCCTTTTGTGGTCTCTGGGGACATTTGGCT 1941
```

```
Qy 101 SerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThr 120
Db 1942 TCTGTTCTAAATGAAGCTGTTTCATCCAGTATTTTCCATCATCTTTGCAAAAATTTATAACC 2001
Qy 121 MetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyPyrSerMetIle 140
Db 2002 ATGTTTGGAAATTAATATAAAACACATTAAGCATGATGCAGAAATTTATTTCATGATA 2061
Qy 141 PheValIleLeuGlyValIleCysPheValSerTyPyrPheMetGlnGlyLeuPheTyGly 160
Db 2062 TTCTGCTATTTGGGTGTTATTTGCTTTGTCAAGTATTTTCATGCAGGGATTTATTTCAGGC 2121
Qy 161 ArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyPyr 180
Db 2122 AGACGAGGGGAAATTTTAAACGATGAGATTAAGACACTTGGCCTTCAAAGCCATGTTATAT 2181
Qy 181 GlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeu 200
Db 2182 CAGGATATTGGCTGGTTGATGAAAGAAACAGCACAGGTTCCAGGATTTGGGCTCTTAACACA 2241
Qy 201 AlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGln 220
Db 2242 GCCATAGATAGCACAAATTCAGGAGCAACAGGTTCCAGGATTTGGGCTCTTAACACA 2301
Qy 221 AsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyPyrGluMetThr 240
Db 2302 AATGCAACTAACATGGGACTTTTCAGTTATCATTTTCCTTTATATATGATGGAGAGATGACA 2361
Qy 241 PheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAla 260
Db 2362 TTCCTGATTTCTGAGTATTTGGCTCCAGTACTTCCCGTGACAGGAATGATTGAAACCGCAGCA 2421
Qy 261 MetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThr 280
Db 2422 ATGACTGATTTGCCAACAAAGATAAGCAAGAACTTAAGCATGCTGGAAAGATAGCAACT 2481
Qy 281 GluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGln 300
Db 2482 GAAGCTTTGGAGAAATATAGCTACTATAGTGTCTATTAAACAAGGAAAGCCCTTCGAGCAA 2541
Qy 301 MetTyGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIle 320
Db 2542 ATGTATGAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCAGAGATTAAT 2601
Qy 321 GlySerCysTyPyrAlaPheSerHisAlaPheIleTyPyrPheAlaTyPyrAlaAlaGlyPheArg 340
Db 2602 GAAAGCTGTTATGATCATTCAGCCATGGCTTTATATATATTTTGGCTATGCAGCAGGGTTTCGA 2661
Qy 341 PheGlyAlaTyPyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPhe 360
Db 2662 TTTGGAGCCTATTTAAATTTCAAGCTGGACGAATGACCCAGAGGGCATGTTTCATAGTTTTT 2721
Qy 361 ThrAlaIleAlaTyPyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyPyr 380
Db 2722 ACTGCAATTTGATATGAGCTATGCGCATCGCAATCGGAAAAACGCTCGTTTTGCTCTCGTAATAT 2781
Qy 381 SerLysAlaLysSerGlyValAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIle 400
Db 2782 TCCAAAGCCAAATCGGGGCTGGCGCATCTGTTTGCCTTTGGTGGAAAGAAACCAANTATA 2841
Qy 401 AspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArg 420
Db 2842 GACAGCCGCACTCAAGAGGGGAAAAAGCCAGACACATGTGAAGGAAATTTAGAGTTTCGA 2901
Qy 421 GluValSerPheTyPyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeu 440
Db 2902 GAAGTCTCTTCTCTATCCATGTCCGACAGATGTTTTTCATCCTCCGTGGCTTATCCCTC 2961
Qy 441 SerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThr 460
Db 2962 AGTATTGAGCGAGGAAAGACAGTAGCATTTTGTGGGGAGCAGCGGCTGTGGGAAAAAGCACT 3021
```



```
QY 461 SerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGly 480
Db 3022 TCTGTTCAACTTCTGAGAGACTTTATGACCCCGTGAAGGACAAGTGTGTTGATGGT 3081
QY 481 ValAspAlaLysGlnLeuLeuValGlnTyrLeuArgSerGlnIleAlaIleValProGln 500
Db 3082 GTGGATGCAAAAGAAATGAATGTACAGTGGCTCGTTCCTCAATAGCAATCGTTCCTCAA 3141
QY 501 GluProValLeuPheAsnGlnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArg 520
Db 3142 GAGCCTGTGCTTCAACTGACGCAATGTGTGAGACATCGCCTATGGTGACACAGCCGT 3201
QY 521 ValValProLeuAspGlnIleLysGluAlaAlaAsnAlaAsnIleHisSerPheIle 540
Db 3202 GTGGTGCCATTAGATGAGATCAAGGAAAGCGCAAAATGCAAGCAATATCCATTCTTTTATT 3261
QY 541 GluGlyLeuProGlnLysTyrAsnThrGlnValGlnValGlnLysGlyAlaGlnLeuSerGly 560
Db 3262 GAAGGTCTCCCTGAGAAATACACACACAAAGTTGGACTGAAAGGACACAGCTTCTGGC 3321
QY 561 GlyGlnGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeu 580
Db 3322 GGCAGAAACAAAGACTAGCTATTGCAAGGGCTCTCTCCAAAACCCAAATTTTATTG 3381
QY 581 LeuAspGluAlaThrSerAlaLeuAspAsnSerGlnLysValValGlnHisAlaLeu 600
Db 3382 TTGGATGAGGCCACTTCAGCCCTCGAATATGACAGTGAGAGGTGGTTGAGCATGCCCTT 3441
QY 601 AspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGln 620
Db 3442 GATAAGCCAGGACGCGAAGGACATGCTTAGTGGTCACTCACAGGCTCTCTGCAATTCAG 3501
QY 621 AsnAlaAspLeuLeuValValLeuHisAsnGlyLysIleLysGlnGlnGlyThrHisGln 640
Db 3502 AACCCAGATTGATAGTGGTTCTGCACATGGAAGATGAAGGAAACCAAGCACTCATCAA 3561
QY 641 GluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 659
Db 3562 GAGCTCCTGAGAAATCGACATATATTTTAAAGTTAGTGAATGCACAGTCAGTGCAG 3618
```

RESULT 5

```
US-09-873-409-13
; Sequence 13, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
; APPLICANT: Frank, Markus
; APPLICANT: Sayegh, Mohamed
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
; TITLE OF INVENTION: Homologue on Chromosome 7p15-21 and Uses Thereof
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873,409
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 13
; LENGTH: 3702
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: No. US20020037522A1e
; LOCATION: (723)..(723)
; OTHER INFORMATION: n at position 723 represents any nucleotide (A, T, C or G)
US-09-873-409-13
```

Alignment Scores:

Pred. No.:	0	Length:	3702
Score:	3323.00	Matches:	659
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	10	Gaps:	0

US-09-873-409-1 (1-659) x US-09-873-409-13 (1-3702)

```
QY 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
Db 1723 ATGTGGCGGAGAAAGGAGCACATGCTGAACATAATGGCAAAACGAGGTCTATATTATTCA 1782
QY 21 LeuValMetSerGlnAspIleLysAlaAspGluGlnMetGluSerMetThrTyrSer 40
Db 1783 CTTGTGATGTACAGATATATAAAAGCTGATGAACACATGGAGTCAATGACATATTCT 1842
QY 41 ThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPhe 60
Db 1843 ACTGAAAGAAAGAACCACTCACTTCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTC 1902
QY 61 IleAspLysAlaGluGlnSerThrGlnSerLysGluIleSerLeuProGluValSerLeu 80
Db 1903 ATTGCAAGGCTGAGAAATCCACCACCAATCTAAAGAGATAAGTCTTCTGAAAGTCTCTCTA 1962
QY 81 LeuLysIleLeuLysLeuAsnLysProGluTyrProPheValValLeuGlyThrLeuAla 100
Db 1963 TTAATAATTTTAAAGTTAAACCAAGCTGAATGCCCTTTTGTGGTCTTCTGGGACATTTGGCT 2022
QY 101 SerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThr 120
Db 2023 TCTGTTCTAAATGGAACTGTTTCATTCAGTATTTTTCATCATCTTTTGCATAAATTTATPAC 2082
QY 121 MetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIle 140
Db 2083 ATGTTTGGAAATTAATGATAAAACCAACATTAAGAGCATGATGCAGAAATTTTATTCATGATA 2142
QY 141 PheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGly 160
Db 2143 TTTCGTCATTTTGGGTGTTTATTTGCTTTTGTGCTTATTTTCATGACGGGATTTATTTACGCG 2202
QY 161 ArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyr 180
Db 2203 AGAGCAGAGGGAATTTTAAACGATGAGATTAAGACATTTGGCCTTCAAGCCATGTATAT 2262
QY 181 GlnAspIleAlaTyrPheAspGlnLysGluAsnSerThrGlyGlyLeuThrThrIleLeu 200
Db 2263 CAGGATATTCCTGCTGTTGATGAAAGGAAACACAGCACAGGAGGCTTGACAAACATATTA 2322
QY 201 AlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGln 220
Db 2323 GCCATAGATATAGCACAAATTAAGGAGCAACAGCTTCAGGATTTGGGCTCTTAACACAA 2382
QY 221 AsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyrGluMetThr 240
Db 2383 AATGCAACTAATACATGGGACTTTTCAGTTATCATTTCTTTATATATGATGGAGATGACA 2442
QY 241 PheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAla 260
Db 2443 TTCCCTGATTTCTGAGTATTGCTCCAGTACTTGCCGTGACAGGAATGATTGAACCCAGCA 2502
QY 261 MetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThr 280
Db 2503 ATGACTGGATTTGCCAACAAAGATAAGCAAGAACTTAAGCATGCTGGAAGATAGCAACT 2562
QY 281 GluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGln 300
Db 2563 GAAGCTTTGGAGATATATAGTACTATAGTGTATTAACAAAGGAAAGAAAGCCCTTCGAGCA 2622
QY 301 MetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIle 320
Db 2623 ATGTATGAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCACAGATTATT 2682
QY 321 GlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArg 340
Db 2683 GGAAGCTTTATGCAATTCAGCCATGCTTTATATATTTTGCCTATGCGACAGGGTTTCCA 2742
QY 341 PheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPhe 360
Db 2743 TTTGGAGCTATTTAATTTCAAGCTGGACGAATGACCCAGAGGGCATGTTCTATAGTTTT 2802
```


Qy 361 ThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyr 380
Db 2803 ACTGCAATTGCATATATGGAGCTATATGCCATCGGAAACGCTGTTTTGGCTCCTGCAATAT 2862
Qy 381 SerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIle 400
Db 2863 TCCAAGCCAATTCGGGGCTCGCATCTGTTTGGCTTTGTTGGAAAAAGAACCAAAATATA 2922
Qy 401 AspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArg 420
Db 2923 GACAGCCGAGTCAAGAGGGAAGGAAAGCCAGACACATGTGAAGGGAATTTAGAGTTTCGA 2982
Qy 421 GluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeu 440
Db 2983 GAAAGTCTCTTTCTATCCATGTCGCCAGATGTTTTTCATCTCCGTGGCTTATCCCTC 3042
Qy 441 SerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThr 460
Db 3043 AGTATTGAGCGAGGAAGACAGTACATTTGTGGGAGCAGCGCTGTGGGAAAGCACT 3102
Qy 461 SerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGly 480
Db 3103 TCTGTCAACTTCTCGACGAGACTTTATGACCCCGTGAAGGACAGTGTCTTTCATGGT 3162
Qy 481 ValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGln 500
Db 3163 GTGGATGCAAAAGAAATTGAATGTACAGTGGCTCCGTTCCCAATAGCAATCGTTCCCTCAA 3222
Qy 501 GluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArg 520
Db 3223 GAGCCTGTGCTCTTCACTGACGACATGCTGAGAACATCGCCTATGGTGACAAACGCGT 3282
Qy 521 ValValProLeuAspGluLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIle 540
Db 3283 GTGGTGCCATAGATGAGATCAAGAGCGGCAATGCAGCAAAATATCCATTTCTTTAT 3342
Qy 541 GluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGly 560
Db 3343 GAAGGTCTCCCTGAGAAATACAAACACAAAGTTGGACTGAAAGGAGCACAGCTTTCTGGC 3402
Qy 561 GlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuGlnLysProLysIleLeuLeu 580
Db 3403 GCCCAGAAACAAAGACTAGCTATTGCAAGGCTCTTCTCCAAAAACCCAAAAATTTATTG 3462
Qy 581 LeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeu 600
Db 3463 TTGGATGAGCCACTTCAGCCCTCGAATATGACGTGAGAGAGTGGTTCAGCATGCCCTT 3522
Qy 601 AspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGln 620
Db 3523 GATAAAGCCAGACGCGGAAGGACATGCCCTAGTGGTCACTCACAGGCTCTCTGCAATTCAG 3582
Qy 621 AsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGln 640
Db 3583 AACGCAGATTTGATAGTGGTCTGCACAATGGAAGATAAAGGAACAAGGAACACTCATCAA 3642
Qy 641 GluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 659
Db 3643 GAGCTCTCTGAGAAATCGACATATATTTTAAGTAGTGAAATGCACAGTCAGTGGCAG 3699

RESULT 6
US-09-880-107-2299
; Sequence 2299, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14

; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2299
; LENGTH: 3924
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 M23234
US-09-880-107-2299

Alignment Scores:
Pred. No.: 5,2e-195 Length: 3924
Score: 1924.50 Matches: 372
Percent Similarity: 74.10% Conservative: 123
Best Local Similarity: 55.69% Mismatches: 162
Query Match: 57.91% Indels: 11
DB: 10 Gaps: 3

US-09-873-409-1 (1-659) x US-09-880-107-2299 (1-3924)

Qy 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
Db 1851 GTAAATTTGGAGCAAGGAAGCCACAGCGAAGCTGATGAAGAAGGAAGGCTGTACTTCAA 1910
Qy 21 LeuVal-----MetSerGlnAspIleLysLysAlaAsp 31
Db 1911 CTTGTCAACATGCAGACATCAGGAAGCCAGATCCAGTCAGAGAATTTGAACATAATGAT 1970
Qy 32 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis 51
Db 1971 GAAAGGCTGCCACTAGAAATGCCCAAAATCGCTGGAATCTCGCCTATTAGGCATTCT 2030
Qy 52 SerValLysSerIleLys---SerAspPheIleAspLysAla---GluGluSerThrGln 59
Db 2031 ACTCAGAAAAACCTTAAATAATTCACAAATGTGTCAAGAGAGCCCTTGATGTGGAACCCGAT 2090
Qy 70 SerLysGluIleSerLeuProGluValSerLeuLysIleLeuLysLeuAsnLysPro 89
Db 2091 GCACTTGAAGCAAAATGTGCCACAGTGTCTTTCTGAAGGCTCTGAAACTGAAATAAACA 2150
Qy 90 GluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPro 109
Db 2151 GAATGGCCCTACTTGTGCTGGAAACAGTATGTGCCATGTCCATGGCAATGGGGGCTTCAGCG 2210
Qy 110 ValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThr 129
Db 2211 GCATTTTCAGTCATATCTCAGAGATCATAGCGATTTTTTGGACCAAGCGATGATGACGTG 2270
Qy 130 LeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPhe 149
Db 2271 AAGCAGCAGAGTGCACATATCTCTTGATTTTCTTATTTCTGGGAATATTTCTTTT 2330
Qy 150 ValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArg 169
Db 2331 TTACTTTCTTCTTCAGGGTTTCAGGTTTGGAAAGCTGGCGAGATCTCCACCAAGAAGA 2390
Qy 170 LeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLys 189
Db 2391 CTGCGGTCAATGGCTTTTAAAGCAATGTCTAAGCAGGACATGAGCTGGTGGTTGATGACCAT 2450
Qy 190 GluAsnSerThrGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGly 209
Db 2451 AAAAAACAGTACTGGTGCACTTTCTACAAGACTTGCACAGATGTGCCCAAGTCCAAAGA 2510
Qy 210 AlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerVal 229
Db 2511 GCCACAGGACCGAGTGGCTTTAATTGACACAGATATAGCTAACTTGGAACTGGTATT 2570
Qy 230 IleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProVal 249


```
Db 2571 ATCATATCATTTATACCGTGGCAGTTAAACCTATTGCTATTAGCAGTTGTTCCAAATT 2630
Qy 250 LeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLys 269
Db 2631 ATTGCTGTGTCAGCAATTTGTAATGAATTTGTTGGCTGGAAATGCCAAAGAGATAAA 2690
Qy 270 GlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleAargThrIle 289
Db 2691 AAAGAACTGGAAGCTGCTGGAAAGATTGCAACAGAGCAATAGAAAAATATTAGGACAGTT 2750
Qy 290 ValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGln 309
Db 2751 GTGTCTTTGACCAGGAAAGAAATTTGAATCAATGTATGTTGAAAAATTTGATGACCT 2810
Qy 310 HisArgAsnThrSerLysLysAlaGlnIleGlySerCysTyrAlaPheSerHisAla 329
Db 2811 TACAGGAATTTCTGTGAGAGGACACATCTATGGAAATTAATCTTTTAGTATCTCACAGCA 2870
Qy 330 PheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGly 349
Db 2871 TTTATGTATTTTCTTCTGCGGTGTTGTTGCGATTGTTGTCATATCTCATTTGGAATGGA 2930
Qy 350 ArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAla 369
Db 2931 CATATCGCTTCAGAGATGTTATTTCTGGTGTCTTCTGCAATTTGTAATTTGTTGTCAGTGGCT 2990
Qy 370 IleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHis 389
Db 2991 CTAGGACATGCCAGTTCATTTGCTCCAGACTATGCTTAAGCTAAGCTGTCTGCAGCCAC 3050
Qy 390 LeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLys 409
Db 3051 TTATTTCATGCTGTTTGAAGACAACTCTGATTGACAGCTACAGTGAAGAGGGGCTGAAG 3110
Qy 410 ProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArg 429
Db 3111 CCTGATAAAATTTGAAGAAATATAACATTTAATGAAGTCGTGTCTCAACTATCCCAACCGA 3170
Qy 430 ProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAla 449
Db 3171 GCNAAGTCGCGAGTCTTACGGGGCTGAGCTTGAGGTGAAGAAAGCCAGACACTAGCC 3230
Qy 450 PheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyr 469
Db 3231 CTGTTGGCAGCAGTGCCTGCTGGAAAGACACGCTGCTCCAGCTCTCTGGAGCGTTCTAC 3290
Qy 470 AspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGln 489
Db 3291 GACCCCTTGGCGGGACAGTGCCTTCTCGATGGTCAAGAAAGCAAAAGAACTCAATGTCCAG 3350
Qy 490 TrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIle 509
Db 3351 TGCTCAGAGCTCAACTCGGAATCGTGTCTCAGAGGCCATCTATTTGACGTGACGAT 3410
Qy 510 AlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGlu 529
Db 3411 GCCGAGAATTTGCCCTATGGAGACACAGCCGGTTGTATCACAGGATGAAATTTGTGAGT 3470
Qy 530 AlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThr 549
Db 3471 GCAGCCAAAAGCTGCCAACATACATCTCTTTCATCGAGACGTTACCCCAACAATATGAACA 3530
Qy 550 GlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAla 569
Db 3531 AGAGTGGGAGATAGGGGACTCAGCTCTCAGGAGGTCAAAAACAGAGGATTTGCTATTGCC 3590
Qy 570 ArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAsp 589
Db 3591 CGAGCCCTCATCAGACAACTCAAACTCTCTCTGTTGATGAGCTACATCAGCTCTGGAT 3650
Qy 590 AsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCys 609
```

```
Db 3651 ACTGAAGTGAAGAGTTGTTCCAGAGCCCTGGACAAAGCCAGAGAGGCCGACCTGC 3710
Qy 610 LeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHis 629
Db 3711 ATTGTGATGTCACCGCTGTCACCATCCAGATGCAGACTTAATAGTGTGTTTCAG 3770
Qy 630 AsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyr 649
Db 3771 AATGGAGAGTCAAGGACATGGCAGCATCAGCAGCTGCTGGCACAAGAAAGCATCTAT 3830
Qy 650 PheLysLeuValAsnAlaGlnSer 657
Db 3831 TTTTCAATGGTCAGTGTCCAGGCT 3854

RESULT 7
US-09-917-800A-1560
; Sequence 1560, Application US/09917800A
; Patent No. US20020119462A1
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Castle, Arthur
; APPLICANT: Elashoff, Michael
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5038-US
; CURRENT APPLICATION NUMBER: US/09/917,800A
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/222,040
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: US 60/222,880
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: US 60/290,029
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/290,645
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: US 60/292,336
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/295,798
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/297,457
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,884
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,459
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 1740
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 1560
; LENGTH: 3912
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020119462A1 NM_012690
US-09-917-800A-1560

Alignment Scores:
Pred. No.: 1,56e-190 Length: 3912
Score: 1882.50 Matches: 363
Percent Similarity: 73.43% Conservative: 129
Best Local Similarity: 54.18% Mismatches: 165
Query Match: 56.65% Indels: 13
DB: 10 Gaps: 2

US-09-873-409-1 (1-659) x US-09-917-800A-1560 (1-3912)
Qy 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrSer 20
Db 1842 GTCATCGTGGAGCAAGGAAGCCACAGCTGAGCTGATAAAGAAAGGAAGGATCTACTTCAGA 1901
Qy 21 LeuValMetSerGln-----AspIleLysLys 29
```


Db 1902 CTTGTTAAATCGACATCGACAGATCGAGAGCCAGATCCTGTCAGAAAGAAATTTGAAGATTGAGCTA 1961
Qy 30 AlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuPro 49
Db 1962 ACTGATGAAAGGCTGCTGGAGGTGTGGCCCCAAATGGCTGGAAAGCACGCATATTAGG 2021
Qy 50 LeuHisSerValLysSerIleLysSerAsp-----PheIleAspLysAlaGluSer 67
Db 2022 AATTCTACGAAGAAAGTCTGAAAGTTTCACGGGGGCATCAAAATAGGCTGGATGGAA 2081
Qy 68 ThrGlnSerLysGluLysSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 87
Db 2082 ACCAATGAACCTTGCAACAGGTGCCACAGGTCTTTCTGAAGTCTTAAGACTGAAT 2141
Qy 88 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107
Db 2142 AAAACAGAGTGGCCCTACTTTGTGGTGGGACACCTGTGGCCATGCCCAACGGGGCCCTC 2201
Qy 108 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLys 127
Db 2202 CAGCGGCATCTCCATCATCTGTCAGAGATCATAGCTATCTTTGGCCCTGGGATGAC 2261
Qy 128 ThrThrLeuLysHisAspAlaGluIleLysSerMetIlePheValIleLeuGlyValIle 147
Db 2262 ACAGTAAAGCAACAGAGGTAACTGTTCTCGTGGTCTTCTTGGGCCCTAGGATCCAC 2321
Qy 148 CysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThr 167
Db 2322 TCCTCTTTACTTTCTCTCTCAGGGTTTCATCTCGGAAGAGCTGGCGAGATCTCCACC 2381
Qy 168 MetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAsp 187
Db 2382 ACAAGGCTCCGGTCCATGCGCTTCAAAGCAATGCTAAGACAGCATAGAGCTGGTTGAC 2441
Qy 188 GluLysGluAsnSerThrGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIle 207
Db 2442 GATCATAAACACAGTACTGGTGGCTCTCTACAAGACTCGCCACAGAGCTCGCGAGTTC 2501
Qy 208 GlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeu 227
Db 2502 CAAGGACCAACAGGACACAGTGGCTTTAATGTGCACAGAACACACGCCAACCTTGGAAAG 2561
Qy 228 SerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuLeuSerIleAla 247
Db 2562 GGTATTATTATATATTATTATTCGGTTGCAACTGCACCTTCTGCTCTATCAGTTGT 2621
Qy 248 ProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLys 267
Db 2622 CCATTCTGCTGTAGCGGAAATTGTTGAAATGTAATGTTGCTGGCAACGCCAAGAGA 2681
Qy 268 AspLysGlnGluLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleAsp 287
Db 2682 GATAAAAGAGATGGAAGCTGCTGGAAGAGATTGCAACAGAGGCATAGAAATATTGCG 2741
Qy 288 ThrIleValSerLeuThrArgLysAlaPheGluGlnMetTyrGluGluMetLeuGln 307
Db 2742 ACTGTTGTATCTTGACCAAGAGAGAAAATTTGAGTCAATGTATGTTGAAAAATTTACAC 2801
Qy 308 ThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSer 327
Db 2802 GGACCTTACAGGAATTCAGTGGGAAGGCTCATCTACGGCATCATCTTTAGCATCTCA 2861
Qy 328 HisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGln 347
Db 2862 CAAGCATTCATGACTTTCTTAATGCTGGCTGCTTTCGATTTGTTCTTACCTCATTTGG 2921
Qy 348 AlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAla 367
Db 2922 AATGGACATCGCTTCAAGATGTCATCTGGTGTCTCAGCAATCGTGTGGTGCA 2981
Qy 368 MetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAla 387
Db 2982 GTGGCTCTAGGACATCCAGCTCATTTGTCTCCAGACTATGCAAAAGCCAAAGCTGTCTGCA 3041

Qy 388 AlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGly 407
Db 3042 GCATACTTATTAGTCTGTTTGAAGACAACCTCTGATTGACAGCTACAGCAGAGAGGA 3101
Qy 408 LysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrPro 427
Db 3102 ATGTGGCCGGATAAGTTTGAAGGAAGCTGACATTCATTAAGTTGTGTTCATATATCCC 3161
Qy 428 CysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThr 447
Db 3162 ACCGGGCCAATGTGCCAGTGTCTCAGGGGCTGACCTCGAGGTGAAGAGGGGAGACC 3221
Qy 448 ValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArg 467
Db 3222 CTGGCCCTGTGGGAGTAGTGGCTGCGGAAGAGACACCGTGTCCAGTGTCTCGAGCGC 3281
Qy 468 LeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsn 487
Db 3282 TTCTACGACCCCATCGCCGGAACAGTGTCTCTCATGTGTGAGGAAGCAAAAGAACTCAAT 3341
Qy 488 ValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCys 507
Db 3342 GTCCAGTGGTCCGAGCTCACTTGGCATTGTGTCCAGAGGCCCATCTCTTTCAGCTGC 3401
Qy 508 SerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIle 527
Db 3402 AGCATCGCCCAAGAACATCGCTACGGAGACAACACGCGTGTCTGTCTCAGGATGAGATT 3461
Qy 528 LysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyr 547
Db 3462 GTGAGGGGGGCCCAAGAGGCCAACATCCACCCCTTCATTGAGACACTGCCCCAAAAGTAT 3521
Qy 548 AsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAla 567
Db 3522 GAAACAGAGTAGGAGACAGGGGACACAGCTCTCTGAGGGCCAGAACACAGAGATTGCT 3581
Qy 568 IleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAla 587
Db 3582 ATCGCCCGAGCCCTCATCAGACAGCTCGGGTCTTACTGTGATGAAGCCACGTCGGCT 3641
Qy 588 LeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArg 607
Db 3642 TTGGACACTGAGAGTGAAGGTGCTGCAGGAAGCGCTGGACAAAGCCAGGAGGCCCGC 3701
Qy 608 ThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValVal 627
Db 3702 ACTCGATTGTGATCGGCACCGCTGTCCACCATCCAGAACCGCAGACTTGATCTGGTG 3761
Qy 628 LeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnLeuLeuLeuArgAsnArgAsp 647
Db 3762 ATCGACAACGGCAAGTCAAGGAGCACGGCACCCACCGACCTGCTGCCCCAGAAAGCC 3821
Qy 648 IleTyrPheLysLeuValAsnAlaGlnSer 657
Db 3822 ATCTATTCTCCATGCTCAACATTCAAGCT 3851

RESULT 8

US-09-769-097-1
; Sequence 1, Application US/09769097
; Patent No. US20020055128A1
; GENERAL INFORMATION:
; APPLICANT: Kimberly Anne Brun
; APPLICANT: Richard James Chenery
; APPLICANT: Harma Ellens
; APPLICANT: John Anthony Feild
; APPLICANT: Lin Yue
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES
; TITLE OF INVENTION: ENCODING RAT MDRIA AND SCREENING METHODS THEREOF
; FILE REFERENCE: GP-50009-C2
; CURRENT APPLICATION NUMBER: US/09/769,097
; CURRENT FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 09/208,809

;; PRIOR FILING DATE: 1998-12-09
;; PRIOR APPLICATION NUMBER: 09/156,800
;; PRIOR FILING DATE: 1999-09-17
;; PRIOR APPLICATION NUMBER: US99/20770
;; PRIOR FILING DATE: 1999-09-10
;; NUMBER OF SEQ ID NOS: 4
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 1
;; LENGTH: 4369
;; TYPE: DNA
;; ORGANISM: RATTUS RATTUS
US-09-769-097-1

Alignment Scores:

Pred. No.: 2,1e-190 Length: 4369
Score: 1882.00 Matches: 362
Percent Similarity: 74.37% Conservatives: 137
Best Local Similarity: 53.95% Mismatches: 158
Query Match: 56.64% Indels: 14
DB: 10 Gaps: 5

US-09-873-409-1 (1-659) x US-09-769-097-1 (1-4369)

QY 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrSer 20
DB 2140 GTCAATGTGGAGCAAGAAATCATGATGAGCTCATGAGAGAGAAAGAAATTTACTTCAAA 2199
QY 21 LeuValMetSerGln-----AspIleLysLysAlaAspGluGlnMetGluSer 36
DB 2200 CTTGTCTATGACTCAGACAGCAGGAAATGAAATGAAATAGGAATGAAGCTTGTGAATCT 2259
QY 37 -----MetThrTyrSerThrGluArgLysThrAsnSerLeu----- 48
DB 2260 AAAGACGGAATTCATATGTGGACATGCTCTCAAAGATTCGGGATCCAGTCTTAATAAGA 2319
QY 49 ProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGlnSerThr 68
DB 2320 AGAAGATCAACTCGCAAAAGCATCGTGGGCCACATGATCAAGACGGGAACTTAGCACC 2379
QY 69 ---GlnSerLysGluLeuSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 87
DB 2380 AAAGAGCTCTGGATGACGACGACGACCTCCAGCTTCTTTTGGCGGATCTCGAAGTTGAAT 2439
QY 88 LysProGluTyrProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107
DB 2440 TCAACTGAATGGCTTATTTTGTGGTGTGTGTTGTTGTCATTAATATGAGGCTTG 2499
QY 108 HisProValPheSerIlelePheAlaLysIleIleThrMetPheGlyAsnAsnAspLys 127
DB 2500 CAACACGATTCCTCCATAATATTTTCAAAGGTTGTAGGGGTTTTTACAAAAAATGACACC 2559
QY 128 ThrThrLeu--LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 146
DB 2560 CCTGAAATCAGCGCAGACAGCAACTGTGTTTCTTTATTTGTTCTGATCTTGGGATC 2619
QY 147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 166
DB 2620 ATCTCTTCATACGTTTCTTCAGGGCTTCACATTTGGCAAGCTGGAGATCTCTC 2679
QY 167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 186
DB 2680 ACCAAGCGATCCGATACATGCTTCAAATCCATGCTGAGACAGGACATAAGCTGGTTT 2739
QY 187 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206
DB 2740 GATGACCTTAATAAACACCAAGAGCGCTGACACACAGGCTTGCCCAATGACGCTGCTCAA 2799
QY 207 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226
DB 2800 GTGAAGGGGCTACAGGCTTAGGCTTGTGTTATTTACCCAGAACATAGCAATCTTGGG 2859
QY 227 LeuSerValIleIleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIle 246

DB 2860 ACAGGCATCATCATCCCTGATCTACGGCTGGCAATTTGACACATTTTACTCTAGCAATT 2919
QY 247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266
DB 2920 GTTCCCATCATGCTATAGCAGGAGTGTGTAATGAAATGTAATGCTTGGACAAGCGGTG 2979
QY 267 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286
DB 2980 AAGATAAGAGGAAGAACTAGAGGTTCTGGGAAGATCGCTACAGAAGCAATTTGAAACATTT 3039
QY 287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluLeuMetLeu 306
DB 3040 CGCACTGCTGCTCTTCTGACTCGGAGCAGAAAGTTTTGAAACTATATATGCCAGAGCTTG 3099
QY 307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326
DB 3100 CAGATACCATACAGAAATGCTTTGAGAAAGCGCAGCTTTTGGGATCATTCTTCTCCTTC 3159
QY 327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 346
DB 3160 ACCCAGGCGCATGATGATTTCTCTCTATGCTGCTGCTTTCGCGTTTGTATGCTTACTTGGTG 3219
QY 347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366
DB 3220 GCACGAGAACTCATGACATTTGAAATGTTCTGTTAGTATTTCTCAGCTATTGCTTTGGT 3279
QY 367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386
DB 3280 GCATGGCAGCTGGGCGAGTCAGTTCATTCGCTCTGACTACGCGAAAGCCAAAGTCTCG 3339
QY 387 AlaAlaHisLeuPheAlaLeuLeuGlyLysLysProAsnIleAspSerArgSerGlnGlu 406
DB 3340 GCATCCCATCATCATCAGGATCATTCAGAAATCCCTGAGATTTGACAGCTACAGCAGGAG 3399
QY 407 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 426
DB 3400 GCGTTGAAAGCCCTATATGTTGGAAGAAATGTGAAATTTAATGAGTCTATGTTCAACT 3459
QY 427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuIleGluArgGlyLys 446
DB 3460 CCCACCCGACCCACATCCAGTGTCTCAGGGCTGAGCCTAGAGGTGAAGAAAGGCGAG 3519
QY 447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 466
DB 3520 ACGCTGGCCCTCGTGGCGCAGTGGCTGCGGAAGAGTACAGTGGTCCAGCTGCTTGAG 3579
QY 467 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 486
DB 3580 CGCTTCTATGACCCCATGCGCGGAAACAGTGTCTTAGATGGCAAGAAATAAAGCACTC 3639
QY 487 AsnValGlnTyrLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506
DB 3640 AATGTCAGTGGCTCCGCGCCCATCTGGGCATTTGTCTCCAGAGCCCATCTCTGTTGAC 3699
QY 507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValProLeuAspGlu 526
DB 3700 TGCACATCGCGCAGCAACATTGCTTACGAGACAACACAGCCGTGCTGCTCCTAAGGAG 3759
QY 527 IleLysGluAlaAlaAsnAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546
DB 3760 ATCTGTAAGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3819
QY 547 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 566
DB 3820 TACAACACAGAGTGGGAGACAAAGGAGCTCAGCTGTGCGGGCGGAGAGCAGCAGGCAATC 3879
QY 567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 586
DB 3880 GCCATCGCGCGCCCTCGTCAGACAGCCTCACAATCTTACTTCTGATGAAGCGCATCA 3939
QY 587 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606
DB 3940 GCTCTGGATACGGAGAGTGAAGAGGCTGTCAGGAAGGCTGTCAGGAAGGCTGTCAGGAAGG 3999


```
Qy 607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 4000 CGCACCTGATGTGATCGCGCACCGCTGCTCCACCATCCAGAACGACAGCTTGATCGTG 4059
Qy 627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlnLeuLeuArgAsnArg 646
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 4060 GTGATTCAGAACGCGCAGGTCAAGGACACGCGCACCCACCGACGAGCTGCTGGCCCGAGAA 4119
Qy 647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 4120 GGCATCTATTCTCGATGCTGAGTGTGAGGCT 4152
RESULT 9
US-09-769-097-3
; Sequence 3, Application US/09769097
; Patent No. US20020055128A1
; GENERAL INFORMATION:
; APPLICANT: Kimberly Anne Brun
; APPLICANT: Richard James Chenery
; APPLICANT: Harma Ellens
; APPLICANT: John Anthony Feild
; APPLICANT: Lin Yue
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES
; FILE REFERENCE: ENCODING RAT MDRIA AND SCREENING METHODS THEREOF
; CURRENT APPLICATION NUMBER: US/09/769,097
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 09/208,809
; PRIOR FILING DATE: 1998-12-09
; PRIOR APPLICATION NUMBER: 09/156,800
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: US99/20770
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 4425
; TYPE: DNA
; ORGANISM: RATTUS RATTUS
US-09-769-097-3
Alignment Scores:
Pred. No.: 2,14e-190 Length: 4425
Score: 1882.00 Matches: 362
Percent Similarity: 74.37% Conservative: 137
Best Local Similarity: 53.95% Mismatches: 158
Query Match: 56.64% Indels: 14
DB: Gaps: 5
US-09-873-409-1 (1-659) x US-09-769-097-3 (1-4425)
Qy 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrSer 20
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2140 GTCAATTGGGAGCAGGAAGAAATCATGATGAGCTCATGAGAGAGAAAGGAATTTACTTCAA 2199
Qy 21 LeuValMetSerGln-----AspIleLysLysAlaAspGluGlnMetGluSer 36
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2200 CTTGTCATGACTCAGACAGCAGGAAATGAATTAATGAAATGAAAGCTGTGAATCT 2259
Qy 37 -----MetThrTyrSerThrGluArgLysThrAsnSerLeu----- 48
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2260 AAAGAYGAAATTGATAATGTGGCATGCTCTTCAAAAGATTCRGGATCCAGTCTAATAAGA 2319
Qy 49 ProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluSerThr 68
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2320 AGAAGATCAACTCGCAAAAGCATCGTGGGCCACATGATCAAGACGGGAACTTAGCACC 2379
Qy 69 ---GlnSerLysGluLeuSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 87
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2380 AAAGAGGCTCTGGATGACAGCATCTCCAGCTCTCTTTGGCGGATCTCTGAAGTTGAAT 2439
Qy 88 LysProGluTyrProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```
Db 2440 TCAACTGAATGGCCTTATTTTGGTGGTGTATTTTGTGCCTAATAATAATGGAGCGTTG 2499
Qy 108 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLys 127
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2500 CAACGACGATCTTCATATATTTTCAAGGTTGTAGGGGTTTTTACAAAAAATCACACC 2559
Qy 128 ThrThrLeu---LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 146
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2560 CCTGAATCCAGCGGACAGACAGCAACTTGTCTTATTTGTTCTCTGATCCTCTGGGATC 2619
Qy 147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 166
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2620 ATCTCTTTTCATTACGTTTTCCTTCAGGGCTTCACATTTGGCAAAGCTGGAGATCTCTC 2679
Qy 167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPhe 186
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2680 ACCAAGCGACTCCGATACATGCTTCAATCCATGCTGAGACAGGACATTAAGCTGGTTT 2739
Qy 187 AspGluLysGluAsnSerThrGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2740 GATGACCTTAAAAACACACACAGGAGCGCTGACCACAGGCTTGCCAATGACGCTGCTCAA 2799
Qy 207 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2800 GTGAAAGGGGCTACAGGCTCTAGGCTTCTGTTATTATCCCAAGACATAGCAAACTCTTGG 2859
Qy 227 LeuSerValIleIleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIle 246
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2860 ACAGGATCATCATATCCCTGATCTACGGCTGGCAATGACACTTTTACTCTTAGCAATT 2919
Qy 247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2920 GTTCCCATCATTTCTATAGCAGGAGTGGTGAATGAAATGTTGTCTGCAAGCGCTG 2979
Qy 267 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2980 AAAGATAAGAGGAAGTCTGGGAGGATCGCTACAGAAAGCAATTTGAAAACCTTT 3039
Qy 287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3040 CGCACTGCTGCTCTTTGACTCGGAGCAGAAAGTTTGAACATATGATGCCAGAGCTTG 3099
Qy 307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3100 CAGATACCATACAGAAATGCTTTGAAGAAAGCGCAGCGCTTTGGGATCACCTTTCTCCTTC 3159
Qy 327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrIleuIle 346
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3160 ACCCAGGCGCATGATGATTTCTCTATGCTGCTTGTCTTCCGGTTTGTATGCTACTTGGTG 3219
Qy 347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3220 GCACAGAACTCATGACATTTGAAATGTTCTGTAGTATTTCTCAGCTATTGTCTTTGGT 3279
Qy 367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3280 GCCATGGCAGTGGGCGAGGTCAGTTTCCTGCTACTACGCGAAAGCCAAAGCTCTCG 3339
Qy 387 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 406
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3340 GCATCCCAATCATCAGGATCATTTGAGAAAAATCCCTGAGATTGACAGTACAGCAGCGAG 3399
Qy 407 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 426
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3400 GGCCTGAAGCCCTAATATGTTGAAGGAAATGTGAATTTAATGGAGTCTATGTTCAACTAT 3459
Qy 427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3460 CCCACCCGACCAACATCCAGTCTTCAGGGGCTGACCTAGAGGTGAAGAAAGGCGAG 3519
Qy 447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuGln 466
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```


Db	2833	ACCAGGCAATGATGATTTTTCCTATGCTGGATGTTTCCGGTTTGAGCGCTACTTGGTG	2892
Qy	347	GLNALaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaIleValGly	366
Db	2893	GCACATAAACTCATGAGCTTCGAGGATGTTCTGTAGTAGTATTTTCAGCTGTGTGCTTGGT	2952
Qy	367	AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly	386
Db	2953	GCCATGGCGTGGGCAAGTCAGTTTCATTTGCTCTGACTATGCCAAAGCCAAATATCA	3012
Qy	387	AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu	406
Db	3013	GCAGCCCATCATCATGATCATTTGAAAAAACCCTTTGATTTGACAGCTACAGCAGGAA	3072
Qy	407	GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr	426
Db	3073	GGCTTAATGCCGMAACATTTGGAGGAATATGTCATTTGGTGAAGTTGTATTCACATAT	3132
Qy	427	ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys	446
Db	3133	CCACCCGACCGGACATCCAGTGCTTCAGGACATGAGCCTGGAGGTGAAGAAGGGCCAG	3192
Qy	447	ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln	466
Db	3193	ACGCTGGCTCTGTGTGGGACAGTAGTGCTGTGGGAAGACACAGTGGTCCAGCTCTGGAG	3252
Qy	467	ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu	486
Db	3253	CGGTTCTACACCCCTTGGCAGGAAGTGCTGCTTGATGGCAAGAAATAAACGCACTG	3312
Qy	487	AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn	506
Db	3313	AATGTTCACTGGCTCCGACACACCTGGGCATCGTGTCCAGGAGGCCATCCTCTTTGAC	3372
Qy	507	CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu	526
Db	3373	TGCAGCATTTGCTCAGAACATTTGCTATGAGACAAACAGCCGGTGGTGTCCACAGAAAGAG	3432
Qy	527	IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys	546
Db	3433	ATCGTAGGGCAGCAAGGAGGCCAACATACATGCTTCATCGAGTCACTGCCTTAATAAA	3492
Qy	547	TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu	566
Db	3493	TATAGCACTAAAGTAGAGACAAAGGAACCTCAGCTCTCTGGTGGCCAGAACACGCATT	3552
Qy	567	AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer	586
Db	3553	GCCATAGCTCGTCCCTTGTTAGACAGCCTCATATTTTGTCTTTGGATGAAGCCACGTC	3612
Qy	587	AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly	606
Db	3613	GCTCTGGATACAGAAAGTGAAGAGTTGTCCAAAGAGCCCTGGCAAGCCAGAGAGGC	3672
Qy	607	ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal	626
Db	3673	CGACCTGCACTTGATTTGCTCACCGCTGTCCACCATCCAGATCCAGACTTAAATAGTG	3732
Qy	627	ValLeuHisAsnGlyLysLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg	646
Db	3733	GTGTTTCAGAATGGCAGAGTCAAGGAGCATGCCAGCATCCAGCAGTCTGTCGCACAGAA	3792
Qy	647	AspIleTyrPheLysLeuValAsnAlaGlnSer	657
Db	3793	GGCATCTAATTTTCAATGGTCAAGTGTCCAGGCT	3825
RESULT 11			
US-09-866-866A-3			
; Sequence 3, Application US/09866866A			
; Patent No. US20020102244A1			
; GENERAL INFORMATION:			
; APPLICANT: Sorrentino, Brian			

Qy 108 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAsp 126
Db 2594 CAACAGCATTTGCAATAATATTTTCAAGATTATATAGGGTTTTACAGAATTTGATGAT 2653
Qy 127 LysThrLeuLysHisAspAlaGluLleIleThrSerMetIlePheValIleLeuGlyVal 146
Db 2654 CCGTGAACCAAAACAGACAGATAGTAACTTGTGTTTCACTATTTGTTTCTGACCCCTGGAAAT 2713
Qy 147 IleCysPheValSerTyrPheMetGlnGlyLeuPheThrGlyArgAlaGluLleLeu 166
Db 2714 ATTTCTTTTATTAATTTTCTTCAGGGTTTCACATTTGGCAAGCTGGAGAGATCTCTC 2773
Qy 167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 186
Db 2774 ACCAAGCGCTCCGATACATGTTTCCGATCCATGCTCAGACAGATGTGAGTTGGTTT 2833
Qy 187 AspGluLysGluAsnSerThrGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206
Db 2834 GATGACCCCTAAAACACCACTGGAGCATTTGACTACAGGCTCGCAATGATGCTGCTCAA 2893
Qy 207 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226
Db 2894 GTTAAAGGGCTATAGTTCAGGCTTGCTGTAATTAACCCAGATATACCAATCTTGGG 2953
Qy 227 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle 246
Db 2954 ACAGAAATATATATCTTCATCTATGTTGGCAACTAACACTGTTACTCTTAGCAATT 3013
Qy 247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266
Db 3014 GTACCATCATTTGCAATAGCAGGAGTTGTTGAATGAAATGTTGTCTGGACAAGCACTG 3073
Qy 267 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286
Db 3074 AAAGATAGAAGAACTAGAGGCTGCGGAAGATGCTACTGGAAGCAATAGAAAACCTTC 3133
Qy 287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluMetLeu 306
Db 3134 CGAACCGTTGTTCTTGTGACTCAGGACAGAGATTGGAACATATGTATGCTCAGAGTTTG 3193
Qy 307 GlnThrGlnHisArgAsnThrSerIysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326
Db 3194 CAGGTACCATACAGAACTCTTTGAGAAAGCACACATCTTTGGAAATTACATTTTCCCTTC 3253
Qy 327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGluAlaTyrLeuIle 346
Db 3254 ACCCAGGCAATGATGATTTTCTCTATGCTGGATGTTCCGGTTTGGAGCCTACTTGGTG 3313
Qy 347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366
Db 3314 GCACATAAATCATGAGCTTTGAGGATGTTCTGTAGTATTTTACGCTGTTGCTTGGT 3373
Qy 367 AlaMetAlaIleGlyThrLeuValLeuAlaProGluTyrSerIysAlaLysSerGly 386
Db 3374 GCCATGGCCGTTGGGCAAGTCACTTCAATTTGCTCTGACTATGATCCCAAGCCAAATATCA 3433
Qy 387 AlaAlaHisLeuPheAlaLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 406
Db 3434 GCAGCCCATCATCATGATGATCATTTGAAAGAAACCCCTTTGATTTGACAGCTACAGCAGGAA 3493
Qy 407 GlyLysLysProAspThrCysGluGlyAsnLeuPheArgGluValSerPhePheTyr 426
Db 3494 GGCCTTAATCCGAACACATTTGAAGGAATATGACATTTGGTGAAGTTGATTTCAACTAT 3553
Qy 427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446
Db 3554 CCCACCCGACCGACATCCAGTCTTCAGGACTTCAGGCTGAGGCTGAAGAGGCCAG 3613
Qy 447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuGln 466
Db 3614 ACGCTGGCTCTGGTGGCAGCAGTGGCTGTGGGAAGAGACACAGTGTGCTCAGCTCTGGAG 3673

Qy 467 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 486
Db 3674 CGGTCTTACGACCCCTTGGCAGGAAAGTGTCTGTATGGCAAGAAATAAAGCACTG 3733
Qy 487 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506
Db 3734 AATGTTCTAGTGGCTCCGAGCACACCTGGGCATCGTGTCCAGAGCCCATCTCTGTTGAC 3793
Qy 507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValProLeuAspGlu 526
Db 3794 TGCAGCATTTGCTGAGAACATTTGCCCTATGGAGACAACAGCCGGTGTGTACAGGAGAG 3853
Qy 527 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546
Db 3854 ATTTGTAGGGGCAAGAGGAGGCCAACATACATGCTCTTCATCGAGTCACTGCTTAATAA 3913
Qy 547 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 566
Db 3914 TATGACATTAAGTAGGAGACAAGAACTCAGCTCTCTGTGGGCCAGANAACAGCATT 3973
Qy 567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuAspGluAlaThrSer 586
Db 3974 GCCATAGCTCGTGCCTTTGTAGACAGCTCATATTTTGTCTTTGGATGAACCCAGTCA 4033
Qy 587 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606
Db 4034 GCTCTGGATACAGAAAGTGAAGGTTGTCCAAGAAAGCCCTGGACAAAGCCAGAGAGGC 4093
Qy 607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626
Db 4094 CGCAGCTGCATTTGATTTGCTCAGCCCTGTCCACCATCCAGAAATGCAGACTTAATAGTG 4153
Qy 627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuArgAsnArg 646
Db 4154 GGTGTTCAAGATGGCAGAGTCAAGAGCATGGCAGCATCAGCAGCTGCTGCGCAGAAA 4213
Qy 647 AspileTyrPheLysLeuValAsnAlaGlnSer 657
Db 4214 GGCATCTATTTTCAATGCTCAGTGTCCAGGCT 4246
RESULT 13
US-09-306-417-1
; Sequence 1, Application US/09306417
; Patent No. US20020103144A1
; GENERAL INFORMATION:
; APPLICANT: Heinrich-Pette-Institut
; TITLE OF INVENTION: Retroviral Gene Transfer Vectors
; FILE REFERENCE: P50491
; CURRENT APPLICATION NUMBER: US/09/306,417
; CURRENT FILING DATE: 1999-05-06
; EARLIER APPLICATION NUMBER: DE 198 22 115
; EARLIER FILING DATE: 1998-05-08
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 1
; LENGTH: 8630
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: proviral
; OTHER INFORMATION: plasmid DNA
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(160)
; OTHER INFORMATION: plasmid backbone (pUC)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (161)..(677)
; OTHER INFORMATION: 5'-LTR
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (532)..(1219)
; FEATURE:

Qy	587	AlaLeuAspAsnAaspSerGluIysValValGlnHisAlaLeuAspLysAlaArgThrGly	606
		::: ::: ::: ::: ::: ::: :::	
Dd	4832	GCTCTGGATACAGAAAGTGAAGAAGTTGTCCAAAGAGCCCTGCACAAAGCCAGAGAAGGC	4891
Qy	607	ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal	626
		::: ::: ::: ::: ::: ::: :::	
Dd	4892	CGCACCTGGCAATTGGATGTCTACC GCCCTGTCCACATTCAGAAATGCCAGACTTAATATGTG	4951
Qy	627	ValLeuHisAsnGluLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg	646
		::: ::: ::: ::: ::: ::: :::	
Dd	4952	GTGTTTCAGAAATGGCAGAGTGTCAGAGAGCATGGCAGCATCAGCAGCTCTCTGGCAGAAA	5011
Qy	647	AspIleTyrPheIysLeuValAsnAlaGlnSer	657
		::: ::: ::: ::: ::: ::: :::	
Dd	5012	GGCATCTATTTTTCAATGGTCAGTGTCCAGGCT	5044

RESULT 14

```

US-09-306-417-2
; Sequence 2, Application US/09306417
; Patent No. US20020103144A1
; GENERAL INFORMATION:
; APPLICANT: Heinrich-Pette-Institut
; TITLE OF INVENTION: Retroviral Gene Transfer Vectors
; FILE OF INVENTION: P50491
; CURRENT APPLICATION NUMBER: US/09/306,417
; CURRENT FILING DATE: 1999-05-06
; EARLIER APPLICATION NUMBER: DE 198 22 115
; EARLIER FILING DATE: 1998-05-08
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 8630
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: proviral
; OTHER INFORMATION: plasmid DNA
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(8630)
; OTHER INFORMATION: retroviral expression vector SPbeta91msA1

```

Alignment Scores:		
Pred. No.:	6.95e-189	Length: 8630
Score:	1872.00	Matches: 356
Percent Similarity:	74.2%	Conservative: 142
Best Local Similarity:	53.0%	Mismatches: 159
Query Match:	56.3%	Indels: 14

DB: 10 Gaps: 4

US-09-873-409-1 (1-659) x US-09-306-417-2 (1-8630)

Qy 1 MetLeuAlaGluLysGlyValahisalaGluLeuMetAlaLysArgGlyLeuTyrSer 20
Db : : : : : | | | | | | | | | | : : : : : | | | | : : : : :
3032 GTCAATGTGGAGAGAAGAAATCATGTAACCTCATGAAGAAGAAGGCATTACTCTCAA 3091
Qy 21 LeuValMetSerGln-----AspIleLysLysAlaAspGluGlnMet 34
Db : : : : : | | | | | : : : : : | | | : : : : :
3092 CTTGTCACAATGCAGACACAGCGAAAATGAAGTTGAATTAGAAAAATGCAGCTGATGAATCC 3151
Qy 35 GluSerMetThrTyrSerThrGluArgLysThrAenSerLeuProLeuHisSerVal--- 53
Db : : : : : | | | | : : : : : | | | : : : : :
3152 AAAAGTCAAAATTGATCCTCTGGAAATGTCTTCAAAATGATTCAAAGATCCAGCTTAATAAGA 3211
Qy 54 -----LysSerIleLys--SeraspPheIleAspLysAlaGluGluSer 67
Db : : : : : | | | | : : : : : | | | : : : : :
3212 AAAGATCAACTCGTAGGAGTGTCTGTGATCACAGCCCNAAGACAGAAAGCTTAGTAGTACC 3271
Qy 68 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 87
Db : : : : : | | | | : : : : : | | | : : : : :
3272 AAAGAGCTCTGGATCAAGATATACCTCCAGTTTCCTTTTGGAGGATTAATGAAGCTAAAT 3331
Qy 88 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107
Db : : : : : | | | | : : : : : | | | : : : : :
3332 TTAACCTGAATGGCTTATTTCTGTGTGTATTTTGTGCCATTATAAATGGAGGCGCT 3391
Qy 108 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAsp 126
Db : : : : : | | | | : : : : : | | | : : : : :
3392 CAACCGACATTTCCAATAATATTTCAAAGATATAGGGGTTTTTACAAGAAATGATGAT 3451
Qy 127 LysThrThrLeuLysHisAspAlaGluIleLysSerMetIlePheValIleLeuGlyVal 146
Db : : : : : | | | | : : : : : | | | : : : : :
3452 CCTCAACAAACACGACAGATAGTAACCTGTGTTTCACATTTGTTCTAGCCCTTGGAAAT 3511
Qy 147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 166
Db : : : : : | | | | : : : : : | | | : : : : :
3512 ATTTCTTTTATTAACATTTTCTTCAAGGTTTCACATTTGGCAAGCTGGAGAGATCCTC 3571
Qy 167 ThrMetArgLeuHrgHisLeualaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 186
Db : : : : : | | | | : : : : : | | | : : : : :
3572 ACCAAGCGGCTCGATACATGTTTTCGATCCATGCTCACACAGAGATGTGAGTTGGTTT 3631
Qy 187 AspGluLysGluAsnSerThrGlyGlyLeuThrThrilLeuAlaIleAspIleAlaGln 206
Db : : : : : | | | | : : : : : | | | : : : : :
3632 GATGACCTCAAAAACACCACTGGAGCATGTACTACCGGCTCGCCAAATGATGCTGCTCAA 3691
Qy 207 IledGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226
Db : : : : : | | | | : : : : : | | | : : : : :
3692 GTTAAAGGGGCTATAGTTTCCAGGCTTGCTGTAATTAATCCAGAAATATAGCAAACTTTGGG 3751
Qy 227 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle 246
Db : : : : : | | | | : : : : : | | | : : : : :
3752 ACAGGAATAATATATCTTCACTATGTTGGCAACTAACACCTGTTACTCTCTTAGCAAAAT 3811
Qy 247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266
Db : : : : : | | | | : : : : : | | | : : : : :
3812 GTACCCATCATTCATACAGAGAGTTGTTGAAATGAATAATGTTGTCTGGNACAGCACTG 3871
Qy 267 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286
Db : : : : : | | | | : : : : : | | | : : : : :
3872 AAAGATAAGAAAGAACTAGNAGGTCCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTC 3931
Qy 287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306
Db : : : : : | | | | : : : : : | | | : : : : :
3932 CGAACCGTGTGTTCTTTTGACTCAGGACGAGAAGTTTGAACATATGTATGCTCAGAGTTTG 3991
Qy 307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyralaPhe 326
Db : : : : : | | | | : : : : : | | | : : : : :
3992 CAGGTACCATCAGAAACCTCTTTGGAGAAAGACACACATCTTTTGGAAATACATTTCTCTC 4051
Qy 327 SerHisAlaPheIleTyrPheAlaTrpAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 346


```
Db 4052 ACCAGCAATGATGATTTTCTATGCTGGAGTGTTCGGTGTGGAGCCTACTTGGTG 4111
Qy 347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaIleGly 366
Db 4112 GCACATAAATCATGAGCTTTGAGGATGTTCTGTAGTATTTTCAGCTGTGTCTTGGT 4171
Qy 367 AlaMetAlaIleGlyThrLeuValLeuAlaProGluTyrSerIysAlaIysSerGly 386
Db 4172 GCCATGCCCTGGGCAAGTCAGTTTCATTTCTGCTCCTGACTATGCCAAGCCAAATATCA 4231
Qy 387 AlaAlaHisLeuPheAlaLeuLeuGluIysLysProAsnIleAspSerArgSerGlnGlu 406
Db 4232 GCAGCCACATCATCATGATCATTTGAAAAACCCCTTTGATTGACAGCTACACACGGAA 4291
Qy 407 GlyLysLysProAspThrCysGluGlyAsnLeuPheArgGluValSerPhePheTyr 426
Db 4292 GGCTAATGCCGACACATTTGGAGGAAATGTCACATTTGGTGAAGTTGATTCAACTAT 4351
Qy 427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446
Db 4352 CCCACCCGACCGGACATCCCGAGTCTTCAGGACTGAGCCTGGAGGTGAAGAAGGCCAG 4411
Qy 447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 466
Db 4412 ACCTGGCTCTGTGGGACAGTGGCTGGGAAGAGCACAGTGGTCCAGCTCCCTGGAG 4471
Qy 467 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 486
Db 4472 CGGTTCTACAGCCCTTGGCAGGAAAGTCTGCTGTGAGGAAAGTGAAGAGGCGACTG 4531
Qy 487 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506
Db 4532 AATGTTTCAGTGGCTCCGAGCACACATCGTGGGATCGTGTCCAGGAGCCATCTGTTTAC 4591
Qy 507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 526
Db 4592 TGACGATTCCTCAGAACATTTGCTTATGGAGAACACAGCCGGTGGTGTCCAGGAAGAG 4651
Qy 527 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546
Db 4652 ATCGTGAGGCGACCAAGGAGGCGCAATACATGCTTCTATCAGTCACTGCCCTAATAAA 4711
Qy 547 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 566
Db 4712 TATAGCACTAAAGTAGAGACAAAGGAATCAGCTCTCTGCTGGCCAGAAACACGCATT 4771
Qy 567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 586
Db 4772 GCATAGCTCGTCCCTTGTAGACAGCCTCATATTTTGTCTTGTGATGAAGCCAGCTCA 4831
Qy 587 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606
Db 4832 GCTCTGATACAGAAAGTGAAGAGTGTTCAGAGAGCCTTCAGAACAGCCAGAGAGGC 4891
Qy 607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626
Db 4892 CGCACCTGCTGATGTTGCTCAGCCGCTGTCCACCATCCAGAAATCCAGACTTAATAGTG 4951
Qy 627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 646
Db 4952 GTGTTTCAGAAATGGCAGAGTCAAGAGGACATGGCCAGCATCAGAGCTGTGGCAGAGAA 5011
Qy 647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657
Db 5012 GGCACTATTTTCAATGGTCACTGTGCCAGGCT 5044
```

RESULT 15

```
US-09-866-866A-7
; Sequence 7, Application US/09866866A
; Patent No. US20020102244A1
; GENERAL INFORMATION:
; APPLICANT: Sorrentino, Brian
```

```
; APPLICANT: Schuetz, John
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
; FILE REFERENCE: 1340-1-021CIP2
; CURRENT APPLICATION NUMBER: US/09/866,866A
; CURRENT FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 09/584,586
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: PCT/US99/11825
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: 60/086,988
; PRIOR FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 4788
; TYPE: DNA
; ORGANISM: Mus musculus
; US-09-866-866A-7
```

```
Alignment Scores:
Pred. No.: 3,6e-189 Length: 4788
Score: 1871.00 Matches: 352
Percent Similarity: 73.70% Conservative: 144
Best Local Similarity: 52.30% Mismatches: 159
Query Match: 56.30% Indels: 18
DB: 10 Gaps: 4
```

```
US-09-873-409-1 (1-659) x US-09-866-866A-7 (1-4788)
Qy 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
Db 1801 GTCATGTGCGACCAAGAAATCATGATGAGTCTCATGAGAGAAAGGGCATTTACTTCAA 1860
Qy 21 LeuValMetSerGln-----AspIleLysLysAla 30
Db 1861 CTGTGTCATGACACAGACAGCAGCAAGAAATGAATTAATAGGAAATGAAGCTGTGAATCT 1920
Qy 31 AspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeu 50
Db 1921 AAGGATGAAATTAATTAATTTAGACATGCTTCAAAAGATTAGGATCCAGTCTAATAAGA 1980
Qy 51 HisSerValLysSerIleLysSerAspPheIle-----AspLysAlaGlu 65
Db 1981 AGA-----AGATCAACTCGCAAAAGACATCTGTGGACCATGACCAAGACAGAGCTT 2034
Qy 66 GluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLysLysIleLeuLys 85
Db 2035 AGTACCAAGAGCCCTGGATGAAGATGATCTCCAGCTTCCTTTGGCGGATCCTGAAG 2094
Qy 86 LeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGly 105
Db 2095 TTGAATTTCAACTGAATGGCTTATTTTGTGGTGTATATTCTGTGCCATAATAAATGGA 2154
Qy 106 ThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn--- 124
Db 2155 GCGTTCACCGCCAGCATCTCCGTAATATTTTCAAAAGTTGTAGGGGTTTTTACAAATGT 2214
Qy 125 AsnAspLysThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeu 144
Db 2215 GGCCCCCTGAAACCCAGCGGAGACAGCAACTGTGTTTCTCTGTGTCTGTATCTT 2274
Qy 145 GlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGlu 164
Db 2275 GGGATCATTTCTTTCATTACATTTTCTTCAGGGCTTCACATTTTGGCAAGCTGGAGAG 2334
Qy 165 IleLeuThrMetArgLeuArgHisIleAlaPheLysAlaMetLeuTyrGlnAspIleAla 184
Db 2335 ATCCCTCACCAAGCGACTCCGATACATGTTTTCAAATCCATCTGACAGCAGGATGTGAGC 2394
Qy 185 TrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrIleLeuAlaIleAspIle 204
Db 2395 TGGTTTGATGACCCATAAAACACCACCGGAGCAGTGCACCCAGGCTCGCCACGATGCT 2454
```



```
Qy 205 AlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsn 224
Db 2455 GCTCAAGTGAAGGGGCTACAGGGTCTAGGCTGTCTGTGATTTTCCAGAAACATAGCAAT 2514
Qy 225 MetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeu 244
Db 2515 CTTGGGACAGGAATCATCATATCCCTATCTATGCTGGCTGGCAACTAACACTTTTACTCTTA 2574
Qy 245 SerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPhe 264
Db 2575 GCAATTGTACCCATCATTCGATAGCTGGAGTGTGAAATGAAATGTTTCTGTCGACAA 2634
Qy 265 AlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGlu 284
Db 2635 GCACTGAAGATAAGAAAGAACTAGAGGTTCTCGAAGGATTGTCTACGGAAGCAATTGAA 2694
Qy 285 AsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGlu 304
Db 2695 AACTTCGCACTGTGTCTCTTTGACTCGGAGCAGAAAGTTTGAAACCATATGTATGCCGAG 2754
Qy 305 MetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyr 324
Db 2755 AGCTTGAGATACCATACAGAAATCGCATGAGAAAGCACAGCTGTTGGGATCACGTTT 2814
Qy 325 AlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyr 344
Db 2815 TCCTTCAACCCAGGCCATGATGATTTTCTTATGCTGCTGCTGTTTCGGTTCGGTGCCTAC 2874
Qy 345 LeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAla 364
Db 2875 TTGGTGACACAAACAACCTACGACTTTTGAAATGTTCTGTAGTATTCTCAGCTATTGTC 2934
Qy 365 TyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLys 384
Db 2935 TTTGGTCATGGCAGTGGGCGAGTCAGTTCATTCGCTCTCTGACTATGCCAAGCAACA 2994
Qy 385 SerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSer 404
Db 2995 GTGTCAAGATCCCATCATCAGGATCATTCAGAAACCCCGAGATTGACAGCTACAGC 3054
Qy 405 GlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhe 424
Db 3055 ACGCAAGGCCCTAAAGCCGAATATGTTGGAAGGAAATGTGCAATTTAGTGGAGTCTGTTC 3114
Qy 425 PheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArg 444
Db 3115 AACTATCCACCCGACCCAGCATCCAGTGTCTCAGGGCTGAGCTTGAGGTGAAGAAG 3174
Qy 445 GlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeu 464
Db 3175 GGCCAGACGCTGGCCCTGTTGGGCGAGCAGTGGCTGGGGAAGAGCACAGTGTCTCCAGCTG 3234
Qy 465 LeuGlnArgLeuTyrAspProValGlnGlnValLeuPheAspGlyValAspAlaLys 484
Db 3235 CTGAGCGCTTCTACGACCCCATGGCTGGATCAGTGTCTTAGATGGCAAAAGATAAAG 3294
Qy 485 GluLeuAsnValGlnTyrLeuArgSerGlnIleAlaIleValProGlnGluProValLeu 504
Db 3295 CAACCTGAATGTCCAGTGGCTCCGAGCAGACGCTGGGCGATTGTCTCCAAAGGCCCAATCTC 3354
Qy 505 PheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeu 524
Db 3355 TTTGACTGCAGGATCGCAGAGAACATTGGCTACGGAGACCAACAGCGGGTCTGTCTTAT 3414
Qy 525 AspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuPro 544
Db 3415 GAGGAGATTGTGAGGCGCAGCCAAAGAGGCGCAACATCCACAGTTTCATCGACTCGCTACCT 3474
Qy 545 GluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGln 564
Db 3475 GATAAATAACACACAGAGTAGGACAAAGGCATCTCAGCTCTCGGGTGGGAGAGCAG 3534
Qy 565 ArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAla 584
```

```
Db 3535 CGCATCCCATCGCACGCGCCCTCGTCAGACAGCCTCACATTTTACTTCTGGACGAAGCA 3594
Qy 585 ThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArg 604
Db 3595 ACATCAGCTCTGGATACAGAAAGTGAAGGTTGTCAGGAAGCGTGGACAAAGCCAGG 3654
Qy 605 ThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeu 624
Db 3655 GAAGGCGCACCTGCATTTGTGATCGCTCCGCGCTGTCACCATCCAGAACCGGACTTG 3714
Qy 625 IleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlnLeuLeuArg 644
Db 3715 ATCGTGTGATTTCAGAACGGCAAGTCAAGAGCACGGCACCCACACGACGCTGCTGCGG 3774
Qy 645 AsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 657
Db 3775 CAGAAGGGCATCTACTTCTCAATGGTCAAGTGTGAGGCT 3813
```

Search completed: March 31, 2003, 15:57:49
Job time : 327.871 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 30, 2003, 01:24:25 ; Search time 8908.97 Seconds
(without alignments)
2152.747 Million cell updates/sec

Title: US-09-873-409-1
Perfect score: 3323
Sequence: 1 MIAEKGHAELMAKRLGLYS.....QELLNRDIYKLVNAQSVQ 659

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2.1/USPTO.spool/US09873409/runat_27032003.115418.19228/app_query.fasta_1.7544
-DB=GenEmbl -QFWT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pcpt -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFWT=ptp -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09873409@cgn.1_54887 @runat_27032003.115418.19228 -NCPU=6 -ICPU=3
-NO XLPXY -NO WMAP -LARGQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:

- 1: gb_ba:*
- 2: gb_htg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sv:*
- 13: gb_un:*
- 14: gb_vi:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_mu:*
- 20: em_om:*
- 21: em_or:*
- 22: em_ov:*
- 23: em_pat:*
- 24: em_ph:*
- 25: em_pl:*
- 26: em_ro:*
- 27: em_sts:*
- 28: em_un:*

- 29: em_vi:*
- 30: em_htg_hum:*
- 31: em_htg_inv:*
- 32: em_htg_other:*
- 33: em_htg_mus:*
- 34: em_htg_pln:*
- 35: em_htg_rod:*
- 36: em_htg_mam:*
- 37: em_htg_vrt:*
- 38: em_sy:*
- 39: em_htgo_hum:*
- 40: em_htgo_mus:*
- 41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	3323	100.0	2066	6	AX339027 Sequence
2	3323	100.0	2856	6	AX339028 Sequence
3	3323	100.0	3177	6	AX339030 Sequence
4	3323	100.0	3621	6	AX339032 Sequence
5	3323	100.0	3702	6	AX339031 Sequence
6	3163	95.2	3699	6	AX478104 Sequence
7	1924.5	57.9	3924	6	AX024455 Sequence
8	1924.5	57.9	3924	6	AX409652 Sequence
9	1924.5	57.9	3924	9	HUMMDR3
10	1916	57.7	3987	10	CRUGP1185
11	1916	57.7	4296	10	CRUGP1
12	1916	57.7	4304	10	CRUGP1165
13	1915.5	57.6	4084	10	MUSMDRA
14	1911	57.5	4018	5	GGA9799
15	1893	57.0	3858	4	OAU78609
16	1892	56.9	4927	10	AF257746
17	1884	56.7	2752	5	AF099732
18	1882.5	56.7	3912	10	RATPLVCO
19	1882.5	56.7	3912	10	AF286167
20	1879	56.5	4323	10	AF286167
21	1876.5	56.5	3905	10	CRUGP111
22	1873	56.4	4186	6	AX108654
23	1873	56.4	4195	6	AX108656
24	1872	56.3	2726	6	I33621
25	1872	56.3	3840	6	AX481416
26	1872	56.3	3860	6	AX322787
27	1872	56.3	3860	6	AX322789
28	1872	56.3	3988	6	AX024454
29	1872	56.3	4192	9	AF016535
30	1872	56.3	4264	6	AR051647
31	1872	56.3	4264	6	AR051650
32	1872	56.3	4646	6	AX336420
33	1872	56.3	4646	6	AX336708
34	1872	56.3	4646	6	AX391099
35	1872	56.3	4646	6	I49610
36	1872	56.3	4646	9	HUMMDR1
37	1872	56.3	4669	6	I08557
38	1872	56.3	6505	6	AR028671
39	1872	56.3	8630	6	AX012320
40	1872	56.3	8630	6	AX012321
41	1872	56.3	9318	6	AR028672
42	1871	56.3	3682	10	MUSPGLY1A
43	1871	56.3	4788	6	AX322793
44	1871	56.3	4924	10	MUSMDR1A
45	1870	56.3	4279	6	AX105057

ALIGNMENTS

RESULT 1

AX339027 AX339027 2066 bp DNA linear PAT 09-JAN-2002
 LOCUS Sequence 9 from Patent WO0194400
 DEFINITION
 ACCESSION AX339027
 VERSION
 KEYWORDS AX339027.1 GI:18129119
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1
 Frank, M.H. and Sayegh, M.H.
 TITLE A gene encoding a multidrug resistance human p-glycoprotein
 JOURNAL homologue on chromosome 7p15-21 and uses thereof
 PATENT: WO 0194400-A 9 13-DEC-2001;
 THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
 FEATURES
 Location/Qualifiers
 1..2066
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 647 a 403 c 470 g 546 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.8e-259 Length: 2066
 Score: 3323.00 Matches: 659
 Percent Similarity: 100.00% Conservativity: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-09-873-409-1 (1-659) x AX339027 (1-2066)
 QY 1 MetLeuAlaGluLysGlyAlaHisAlaGluMetAlaLysArgGlyLeuTyrTyrSer 20
 DB 87 ATGCTGGCGGAGAAAGGAGACATGCTGAACATAATGGCAAAACAGGCTATATATTC 146
 QY 21 LeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSer 40
 DB 147 CTTGTGATGTCCACAGGATATTTAAAGGAGCTGATGAACAGATGGAGTCAATGACATATCT 206
 QY 41 ThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPhe 60
 DB 207 ACTGAAAGAAAGACCACTCACTCTCTGCACTCTGTGAAGAGATAGTCTTCTGAAGTCTCTTA 266
 QY 61 IleAspLysAlaGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeu 80
 DB 267 ATTGACAAGGCTGAGGAATCCACCAATCTAAAGAGATAGTCTTCTGAAGTCTCTTA 326
 QY 81 LeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAla 100
 DB 327 TTAATAAATTTTAAAGTTAAACAGCCCTGAATGGCTTTTGTGGTCTCGGGGACATGGCT 366
 QY 101 SerValLeuAsnGlyThrValHisProValPheSerIleLeuPheAlaLysIleThr 120
 DB 387 TCTGTTCTAAATGAGACTGTTCATCCAGTATTTTCCATCATCTTTGCAAAATTTATAACC 446
 QY 121 MetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIle 140
 DB 447 ATGTTTGGAAATAATGATAAAACACATTAAGCATGATGATGATGATGATGATGATGATGAT 506
 QY 141 PheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGly 160
 DB 507 TTCGTCAATTTGGGTGTTATTTGCTTTGTGAGTATTTTCATGAGGAGATTTATTTTACGGC 566
 QY 161 ArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyr 180
 DB 567 AGAGCAGGGGAAATTTAAACATGAGATTAAGACACTTGGCCCTTCAAGCCATGTTATAT 626
 QY 181 GlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrIleLeu 200
 DB 627 CAGGATATTTGCTGTTTGTATGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 686

QY 201 AlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGln 220
 DB 687 GCCATAGATATAGACAAATTCAGAGGAGCAACAGGTTCCAGANTGGCTTTTACACAA 746
 QY 221 AsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThr 240
 DB 747 AATGCAACTAATCATGGGACTTTCAGTTATCATTTCTTTATATATGATGGAGATGACA 806
 QY 241 PheIleIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAla 260
 DB 807 TTCCTGATTTCTGAGTATGTCTCCAGTACTTCCCGTGACAGGAATGATTGAAACCGCAGCA 866
 QY 261 MetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThr 280
 DB 867 ATGACTGATTTGCCACAAAGATTAACAGAACTTAAGCATGCTGGAAGATAGCAACT 926
 QY 281 GluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGln 300
 DB 927 GAAGCTTTGGAGAAATATACGTACTATAGTGTCTATTAACAAGGGAAGAAAGCTTCGAGCAA 986
 QY 301 MetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIle 320
 DB 987 ATGTATGAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCACAGATTATT 1046
 QY 321 GlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArg 340
 DB 1047 GGAAGCTGTTATGCATTCAGCCATGCTTTATATATTTTCCCTATGCGCAGGCTTCGA 1106
 QY 341 PheGlyAlaTyrIleLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPhe 360
 DB 1107 TTTGGAGCTTATTTAATTCAGCTGGAGCAATGACCCAGAGGGCATGTTTCATAGTTTTT 1166
 QY 361 ThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyr 380
 DB 1167 ACTGCAATTCATATGAGCTATGGCCATCGGAAACCGCTCGTTTGGCTCCTGATAT 1226
 QY 381 SerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLysLysProAsnIle 400
 DB 1227 TCCAAAGCCAAATCGGGGCTGCGCATCTGTTTGGCTTGTGGAAAGAAACCAATATA 1286
 QY 401 AspSerArgSerGlnGluLysLysProAspThrCysGluGlyAsnLeuGluPheArg 420
 DB 1287 GACAGCGCAGTCAAGAGGAGGAAAGCCAGACACATGTGAAGGAAATTTAGAGTTTCA 1346
 QY 421 GluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeu 440
 DB 1347 GAAGTCTCTTTCTTCATTCATGTCGCCAGATGTTTTCATCTCCGCTGCTTATCCCTC 1406
 QY 441 SerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThr 460
 DB 1407 AGTATTGAGCGAGGAAAGACAGTAGCATTTTGGGGAGCAGCGGCTGTGGAAAGACACT 1466
 QY 461 SerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlnValLeuPheAspGly 480
 DB 1467 TCTGTTTCAACTTCTGACAGACTTTATGACCCCGTCAAGGAGCAAGTGTGTTGATGGT 1526
 QY 481 ValAspAlaLysGluLeuAsnValGlnTyrLeuArgSerGlnIleAlaIleValProGln 500
 DB 1527 GTGGATGCAAAAGATTTGAATGTACAGTGGCTCCGTTCCCAATAGCAATCGTTCTCTCAA 1586
 QY 501 GluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArg 520
 DB 1587 GAGCCTGTGCTCTTCAACTGACAGCTTGTGAGACATCGCTATGTTGTCACACACCGCT 1646
 QY 521 ValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIle 540
 DB 1647 GTGGTGGCCATTAGTAGAGATCAAGAGCCGCAATGTCAGCAAAATATTCATTTCTTTATT 1706
 QY 541 GluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyValGlnLeuSerGly 560
 DB 1707 GAAGGTCTCCCTGAGAAATACACACACAGTTGGACTGAAAGGAGCAGACGCTTCTGGC 1766
 QY 561 GlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeu 580


```
Db 1767 GGCAGAACAAAGACTAGCTATTGCAAGGGCTCTTCCAAAAACCCAAATTTTATTG 1826
Qy 581 LeuAspGluAlaThrSerAlaLeuAspAsnSerGluLysValValGlnHisAlaLeu 600
Db 1827 TTGGATGAGGCACCTTCAGCCCTCGATATGACATGAGAGGTGGTTCAGCATGCCCTT 1886
Qy 601 AspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaLeuGln 620
Db 1887 GATAAAGCCAGGACGGGAAGGACATGCTCTAGTGGTCACTCACAGGCTCTCTGCAATTG 1946
Qy 621 AsnAlaAspLeuValValLeuHisGlnGlyLysAlaLeuLysGluGlnGlyThrHisGln 640
Db 1947 AACGCAGATTTGATAGTGGTCTGCACAAATGGAAGATGAAGAAACAAGGAACATCA 2006
Qy 641 GluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 659
Db 2007 GAGCTCTGAGAAATCGAGACATATATTTTAAGTTAGTGAATGCACAGTCAGTGCAG 2063

RESULT 2
AX339028 2856 bp DNA linear PAT 09-JAN-2002
LOCUS AX339028
DEFINITION Sequence 10 from Patent WO0194400.
ACCESSION AX339028
VERSION AX339028.1 GI:18129120
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Frank, M.H. and Sayegh, M.H.
TITLE A gene encoding a multidrug resistance human p-glycoprotein
homologue on chromosome 7p15-21 and uses thereof
JOURNAL Patent: WO 0194400-A 10 13-DEC-2001.
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES
source
1. .2856
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 888 a 540 c 652 g 776 t
ORIGIN

Alignment Scores:
Pred. No.: 2,77e-259 Length: 2856
Score: 3323.00 Matches: 659
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-873-409-1 (1-659) x AX339028 (1-2856)
Qy 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
Db 877 ATGCTGGCGGAGAAAGGAGCACATGCTGAACATAATGGCAAAACGAGGTCTATATTATTC 936
Qy 21 LeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSer 40
Db 937 CTGTGTGATGTCAGGATATTAATAAAGCTGTATGAACAGATGGAGTCAATGACATATTCT 996
Qy 41 ThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPhe 60
Db 997 ACTGAAAGAAAGACCACTCACTTCTCTGCACCTCTGTGAAGCATCAAGTCAGACTTC 1056
Qy 61 IleAspLysAlaGluSerThrGlnSerLysGluLysSerLeuProGluValSerLeu 80
Db 1057 ATTGCAAGGCTGAGGAATCCCAATCTAAAGAGATAAGTCTTCTCGAAGTCTCTCTA 1116
Qy 81 LeuLysIleLeuLysLeuAsnLysProGluTyrProPheValValLeuLysThrLeuAla 100
Db 1117 TTAATAAATTTTAAGTTAAACAAGCCGTGAATGCCCTTTTGTGGTCTTGGGGACATTTGGCT 1176
```

```
Qy 101 SerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThr 120
Db 1177 TCTGTCTTAAATGGAACCTGTTTCATCCAGTATTTTCCATCATCTTTGCAAAATTTATAACC 1236
Qy 121 MetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIle 140
Db 1237 ATGTTTGGAAATATATGATAAAACACACATTAAAGCATGATGCAGAAATTTATTCATGATA 1296
Qy 141 PheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGly 160
Db 1297 TTCGTCTTTTGGGTGTTTATTTGCTTGTCTAGTTATTTTCATGCAGGGATTTATTTACGGC 1356
Qy 161 ArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyr 180
Db 1357 AGAGCAGGGGAAATTTTAAACGATGAGATTAAGACACTTTGGCCCTTCAAGACCATGTTAT 1416
Qy 181 GlnAspIleAlaTyrPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeu 200
Db 1417 CAGGATATTGCTGTTTGTATGAAGAAACACAGCAGGAGGCTTGACAAATATTA 1476
Qy 201 AlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGln 220
Db 1477 GCCATAGATATAGCACAAATTTCAAGGACCAACAGGTTCCAGGATTTGGCGTCTTAACACAA 1536
Qy 221 AsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyrGluMetThr 240
Db 1537 AATGCAACTAACTGGGACTTTTCAGTTATCATTTCTCTTTATATATGATGGAGATGACA 1596
Qy 241 PheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAla 260
Db 1597 TTCCGTGATTCGATGTTTCTCCAGTACTTCCGTCGACAGGAATGATTTGAACCCGACGA 1656
Qy 261 MetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThr 280
Db 1657 ATGACTGGATTTGCCAACAAAGATAAGCAAGAACTTAAAGCATGCTGGAAGATAGCAACT 1716
Qy 281 GluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGln 300
Db 1717 GAAAGCTTTGGAGAAATATATACGTACTATAGTGTCTATTAACAAGGGAAAAAGCCCTTCGAGCAA 1776
Qy 301 MetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIle 320
Db 1777 ATGTATGAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCACAGATTTAT 1836
Qy 321 GlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArg 340
Db 1837 GGAAGCTGTTATGCAATTCAGCCATGCCCTTTATATATTTTGGCTATGCAGAGGTTTGA 1896
Qy 341 PheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPhe 360
Db 1897 TTTTGGAGCCTATTTAATTCAGCTGGACGAATGACCCAGAGGGCATGTTTCATAGTTTTT 1956
Qy 361 ThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyr 380
Db 1957 ACTGCAATTCATATATGAGACTATGCCCATCGAAAAACGCTGTTTGGCTCCTCAATAT 2016
Qy 381 SerLysAlaLysSerGlyValAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIle 400
Db 2017 TCCAAGGCCAAATCGGGGCTCGGCATCTGTTTGGCTTGTGGAAAAAGAACCAATATA 2076
Qy 401 AspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArg 420
Db 2077 GACAGCCGAGTCAAGAAAGGGAAAAAGCCAGACACATGTGAAGGGAATTTAGAGTTTGA 2136
Qy 421 GluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeu 440
Db 2137 GAAGTCTCTTTCTTATCCATGTCGCCAGATGTTTTCATCCTCCGTGGCTTATCCCTC 2196
Qy 441 SerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThr 460
Db 2197 AGTATTGAGCAGGAAAGACAGTACATTTTGTGGGAGCAGCGGCTGTGGGAAAGCACT 2256
Qy 461 SerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyValLeuPheAspGly 480
```



```

|||||
Db 2257 TCTGTTCAACTTCTGCAGAGACTTTATGACCCCGTGCAGGACAAGTGTGTTGATGCT 2316
Qy 481 ValAspAlaLysGluLeuLeuValGlnTTPLeuArgSerGlnIleAlaIleValProGln 500
Db 2317 GTGGATGCAAAAGAAATGAATGACAGTGTCCGTTCCCAATAGCAATGCTTCTCAA 2376
Qy 501 GluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArg 520
Db 2377 GAGCCTGTGCTCTTCAACTGCACCAITTCGTGAGAACATCGCCTATGTTGACACACCGT 2436
Qy 521 ValValProLeuAspGluLeuLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIle 540
Db 2437 GTGGTGCCATTAGATGAGATCAAGAAGCGCAATATGCAGCAATATCCATTCTTTATT 2496
Qy 541 GluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGly 560
Db 2497 GAAGGTCTCCCTGAGAAATACACACACAGTTGGACTGAAAGAGACACAGCTTCTGGC 2556
Qy 561 GlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuGlnLysProLysIleLeuLeu 580
Db 2557 GGCAGAAACAAAGACTAGCTATTGCAAGGGCTCTTCTCCAAAACCCAAATTTTATTG 2616
Qy 581 LeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeu 600
Db 2617 TTGGATGAGGCCACTTTCAGCCCTCGATAATGACAGTGAGAGGTGGTTTCAGCATGCC 2676
Qy 601 AspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGln 620
Db 2677 GATAAAGCCAGGACGGAGGACATCCCTAGTGGTCACTCACAGGCTCTCTGCAATTCAG 2736
Qy 621 AsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGln 640
Db 2737 AACGCAGATTGATAGTGGTCTGCAATGGAAGATAAGATAAGGAACCAAGAACTCATCA 2796
Qy 641 GluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 659
Db 2797 GAGCTCTGAGAAATCGAGACATATATTTAAAGTTAGTGAATGCACAGTCAGTGCAG 2853

RESULT 3
LOCUS AX339030 3177 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 12 from Patent WO0194400.
ACCESSION AX339030
VERSION AX339030.1 GI:18129122
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE Frank, M. H. and Savegh, M. H.
AUTHORS A gene encoding a multidrug resistance human p-glycoprotein
TITLE homologue on chromosome 7p15-21 and uses thereof
JOURNAL Patent: WO 0194400-A 12 13-DEC-2001;
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES
Location/Qualifiers
Source 1. 3177
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 970 a 601 c 735 g 870 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 3,2e-259 Length: 3177
Score: 3233.00 Matches: 659
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-873-409-1 (1-659) x AX339030 (1-3177)

```

```

Qy 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
Db 1198 ATGCTGGCGGAGAAAGGAGCACATGCTGAACCTAATGGCAAAACGAGGTCTATATTATCCA 1257
Qy 21 LeuValMetSerGlnAspIleLysAlaAspGluGlnMetGluSerMetThrTyrSer 40
Db 1258 CTTGTGATGTCACAGATATTTAAAGCTGATGAACAGATGGAGTCAATGACATATTTCT 1317
Qy 41 ThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPhe 60
Db 1318 ACTGAAGAAAGAACCACTCACTCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTC 1377
Qy 61 IleAspLysAlaGluSerThrGlnSerLysGluLysSerLeuProGluValSerLeu 80
Db 1378 ATTGAAGAGCTGAGAAATCCACCAATCTAAAGAGATAAGTCTTCTCGAAGCTCTCTCA 1437
Qy 81 LeuLysIleLeuLysLeuAsnLysProGluTyrProPheValValLeuGlyThrLeuAla 100
Db 1438 TTAAGAAATTTTAAAGTTAAACCAAGCTGAATGGCTTTTGTGGTCTTGGGGACATTTGGCT 1497
Qy 101 SerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThr 120
Db 1498 TCTGTCTAAATGGAACCTGTTTCATCCAGTATTTTCCATCATCTTTGCAAAATATTAACC 1557
Qy 121 MetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIle 140
Db 1558 ATGTTTGGAAATTAATGATAAAACCAATTAAGCATGATGCAGAAATTTATTTCCATGATA 1617
Qy 141 PheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGly 160
Db 1618 TTTCGTCATTTTCGGGTGTTATTTGCTTGTTCAGTTATTTTCATGCGGATTTATTACGGC 1677
Qy 161 ArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyr 180
Db 1678 AGAGCAGGGGAAATTTAACGATGAGATTAGACACTTGGCCTTCAAGCCATGTTATAT 1737
Qy 181 GlnAspIleAlaTyrPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrLeuLeu 200
Db 1738 CAGGATATTCCTGCTGTTGATGAAAGGAAACACACAGGAGGCTTGACAAACATATTA 1797
Qy 201 AlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGln 220
Db 1798 GCCATAGATATAGCAAAATTCAGGAGCAACAGTTCAGGATTTGGGCTCTTTACACAA 1857
Qy 221 AsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyrGluMetThr 240
Db 1858 AATGCACTAACATGGGACTTTCAGTTATCTTCTTATATATGATGGAGGATGACA 1917
Qy 241 PheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAla 260
Db 1918 TTCTCTGATTTCTGATTTGCTCCAGTACTTGCCTGACAGGAATGATTGAAACCGCAGCA 1977
Qy 261 MetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThr 280
Db 1978 ATGACTGGATTTGCCAACAAAGATAAGCAAGAACTTAAGCATGCTGGAAAGATAGCACT 2037
Qy 281 GluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGln 300
Db 2038 GAAGCTTTGGAGAAATATAGTACTATAGTGTCTATTAAACAGGAGAAAGCCCTTCGAGCAA 2097
Qy 301 MetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIle 320
Db 2098 ATGTATGAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCACAGATATT 2157
Qy 321 GlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArg 340
Db 2158 GGAAGCTGTATGCAATCAGCCATGCTTATATATTTTGGCTATGTCAGCAGGGTTTCCA 2217
Qy 341 PheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPhe 360
Db 2218 TTTGGAGCCTATTTAATTCAGCTGGACGATGACCCAGAGGGCATGTTTCATAGTTTTT 2277
Qy 361 ThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyr 380

```


Db 2278 ACTGCAATTCATATGGAGCTATGGCCATCGAAAAACGCTCGTTTGGCTCCCTGAATAT 2337
Qy SerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIle 400
Db 2338 TCCAAAGCCAAATCGGGGCTGGCATCTGTTGGCTTGTGGAAAAAGAACCAATATA 2397
Qy AspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArg 420
Db 2398 GACAGCCGAGTCAGAAAGGAAAAAGCCAGACACATGTAAGGGAATTTAGAGTTTCGA 2457
Qy GluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeu 440
Db 2458 GAAGTCTCTTCTTATCCATGTCGCCAGATGTTTTCATCCTCCGCGGTCTATCCCTC 2517
Qy SerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThr 460
Db 2518 AGTATTAGCGAGGAAAGACAGTACATTTGTGGGAGCAGCGGTGTGGGAAAGCACT 2577
Qy SerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlnValLeuPheAspGly 480
Db 2578 TCTGTTCAACTTCGACAGACTTTATGACCCCGTGAAGGACAAAGTCTGTTTATGGT 2637
Qy ValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGln 500
Db 2638 GTGGATGCAAAAGAAATTGAATGTACAGTGGCTCCGTTCCCAATAGCAATCGTTCCCTCA 2697
Qy GluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArg 520
Db 2698 GAGCTGTGCTCTTCACTGCAGCAATCTGAGAACATCGCTATGGTGACAAACAGCCGT 2757
Qy ValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIle 540
Db 2758 GTGGTGCCATTAGATGAGATCAAGAAGCCGCAATGCAGCAAAATATCCATTCTTTATT 2817
Qy GluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGly 560
Db 2818 GAAGTCTCTCCCTGAGAAATACAAACACAAAGTTGGACTGAAAGGAGCACAGCTTTCTGC 2877
Qy GlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeu 580
Db 2878 GCCAGAAACAAAGACTAGCTATTGCAAGGGCTCTTCTCCAAAAACCCAAATTTTATG 2937
Qy LeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValGlnHisAlaLeu 600
Db 2938 TTGGATGAGCCACTTCAGCCCTCGATAATGACAGTGAAGAGTGGTTCAGCATGCCCTT 2997
Qy AspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGln 620
Db 2998 GATAAAGCCAGCAGCGGAAGGACATGCCCTAGTGGTCACTCACAGGCTCTCTGCAATTG 3057
Qy AsnAlaAspLeuIleValValLeuHisGlnGlyLysIleLysGluGlnGlyThrHisGln 640
Db 3058 AACGAGATTTCATAGTGGTCTGACCAATGGAAGAGATAAGAAACAAAGGAACACTCATCA 3117
Qy GluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 659
Db 3118 GAGCTCTCGAAGAAATCGACATATATTTAAGTTAGTGAATGCACAGTCACTGCGAG 3174

RESULT 4
AX339032 3621 bp DNA linear PAT 09-JAN-2002
LOCUS Sequence 14 from Patent WO0194400.
DEFINITION AX339032
ACCESSION AX339032
VERSION AX339032.1 GI:18129124
KEYWORDS human.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Frank, M.H. and Sayegh, M.H.
TITLE A gene encoding a multidrug resistance human p-glycoprotein

homologue on chromosome 7p15-21 and uses thereof
Patent: WO 0194400-A 14 13-DEC-2001;
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES
source Location/Qualifiers
1. .3621 /organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 1081 a 696 c 842 g 1002 t
ORIGIN
Alignment Scores:
Pred. No.: 3,81e-259 Length: 3621
Score: 3323.00 Matches: 659
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-873-409-1 (1-659) x AX339032 (1-3621)

Qy 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
Db 1642 ATGCTGGCGGAGAAAGGAGCACATGCTGAACATAATGGCAAAACGAGGTCTATATTATCA 1701
Qy 21 LeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSer 40
Db 1702 CTTGTGATGTCAACAGGATATTAATAAAGCTGATGAACAGATGGAGTCAATGACATATTTCT 1761
Qy 41 ThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPhe 60
Db 1762 ACTGAAGAAAGACCAACTCACTTCCTCTGCACCTCTGTGAAGAGCATCAAGTCAGACTTC 1821
Qy 61 IleAspLysAlaGluGluSerThrGlnSerLysGluLysSerLeuProGluValSerLeu 80
Db 1822 ATTGCAAGGCTGAGGAATCCACCCCAATCTAAGAGATAAGTCTTCTGGAAGTCTCTCTA 1881
Qy 81 LeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAla 100
Db 1882 TTAATAATTTTAAAGTTAAACAAGCCTGAATGGCCTTTTGTGGTCTCTGGGACATTTGGCT 1941
Qy 101 SerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThr 120
Db 1942 TCTGTCTTAAATGGAACTGTTCATCCAGTATTTTCCATCACTTTTGCAGAAATTTATAACC 2001
Qy 121 MetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIle 140
Db 2002 ATGTTTGAATAATATATAAACCACATTAAAGCATGATGCAGAAATTTATTTCCATGATA 2061
Qy 141 PheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGly 160
Db 2062 TTCTGTCAATTTTGGGTGTTTATTTTGTCTAGTTATTTTTCATGTCAGGGGATTTATTTACGGC 2121
Qy 161 ArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheAlaMetLeuTyr 180
Db 2122 AGAGCAGGGGAAATTTTAAACCATGAGATTAAGACACTTGGCCCTTCAAAGCCATGTTATAT 2181
Qy 181 GlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeu 200
Db 2182 CAGGATATTGCTGGTTTGTATGAAAGGAAACAGCACAGGAGGCTTGCAACAATATA 2241
Qy 201 AlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGln 220
Db 2242 GCCATAGATATAGCACAAATTTCAAGGAGCAACAGGTTCCAGGATTTGGCGTCTTAACACAA 2301
Qy 221 AsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThr 240
Db 2302 AATGCAACTAACTGGGACTTTTCAGTTATCATTTTCTTTATATATATGAGTGGAGATGACA 2361
Qy 241 PheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAla 260
Db 2362 TTCTGTATTCGAGTATTTGCTCCAGTACTTGCCTGCAGCAGGAATGATTGAAACCCGACGA 2421
Qy 261 MetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThr 280

Db 2422 ATGACTGGATTTCCTCAACAAAGATAGCAAGAACTTAAGCATGCTGGAAAGATAGCAACT 2481
Qy 281 GluAlaLeuGluAenIleArGThrIleValSerLeuThrArgGluLysAlaPheGluGln 300
Db 2482 GAAGCTTTGAGAAATATACGTACTATAGTGTCTATACAAAGGAAAAAGCCTTCGAGCAA 2541
Qy 301 MetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIle 320
Db 2542 ATGTATGAAGAGATGCTTTCAGACTCAACACAGAAATACCTCGAAGAAGACACAGATTATT 2601
Qy 321 GlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArg 340
Db 2602 GGAAGCTGTATGCAATTCAGCCATGCTTTATATATTTTCCCTATGCAGAGGGTTTCGA 2661
Qy 341 PheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPhe 360
Db 2662 TTTGGAGCCTATTATTTCAAGCTGGAGCAATGACCCAGAGGGCATGTTTCATAGTTT 2721
Qy 361 ThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyr 380
Db 2722 ACTGCAATTCATATGAGCTATGGCCATCGAAAAACGCTCGTTTGGCTCCTGAATAT 2781
Qy 381 SerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIle 400
Db 2782 TCCAAAGCCAAATCGGGGGCTGGCGCATCTGTTGCCTTGTGGAAAGAAACCAATATA 2841
Qy 401 AspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArg 420
Db 2842 GACAGCCGAGTCAGAGGGAAGAAAGCCAGACACATGTGAAGGGAATTTAGAGTTTCGA 2901
Qy 421 GluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeu 440
Db 2902 GAAGTCTCTTCTTCTATCCATGTCGCCAGATGTTTCATCTCCGTGCTTATCCCTC 2961
Qy 441 SerIleGluArgGlyLysThrValAlaPheValIleSerSerGlyCysGlyLysSerThr 460
Db 2962 AGTATTGAGCGAGGAAGACAGTAGCATTTTGTGGGAGACGGCTGTGGAAAGACACT 3021
Qy 461 SerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGly 480
Db 3022 TCTGTTCAACTTCTGCAGACACTTATGACCCCGTGCAAGACAAAGTCTGTTGATGGT 3081
Qy 481 ValAspAlaLysGluLeuAsnValGlnThrLeuArgSerGlnIleAlaIleValProGln 500
Db 3082 GTGGATGCAAAAGAAATTAATGTACAGTGGCTCGCTTCCCAATAGCAATCGTTCCTCAA 3141
Qy 501 GluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArg 520
Db 3142 GAGCCTGTGCTTCTCAACTCGACATTTGCTGAGAACATCGCCCTATGGTGACACAGCCGT 3201
Qy 521 ValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIle 540
Db 3202 GTGGTGCCATTAGATGAGATCAAGAGCGCAATGCAAGCAATATCCATCTCTTTATT 3261
Qy 541 GluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGly 560
Db 3262 GAAGGTCTCCCTGAGAAATCAACACACAAAGTTGGACTGAAGAGCACACAGCTTCTGGC 3321
Qy 561 GlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeu 580
Db 3322 GCCCAGAAACAAAGACTAGCTATTGCAAGGGCTCTCTCCAAAGCCCAAAATTTTATTG 3381
Qy 581 LeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeu 600
Db 3382 TTGATGAGGCCACTTCAGCCCTCGATATACAGTGAAGAGTGGTTTCAGCATGGCCCT 3441
Qy 601 AspLysAlaArgThrGlyArgThrCysLeuValThrHisArgLeuSerAlaIleGln 620
Db 3442 GATAAAGCCAGGACGGAGGACATGCTAGTGGTCACTCACAGGCTCTCTGCAATTCAG 3501
Qy 621 AsnAlaAspLeuIleValValLeuHisAsnGlyValIleLysGluGlnGlyThrHisGln 640

Db 3502 AACGCAGATTTGATAGTGTCTGCACATGGAAGATAAAGAACAAAGAACTCATCA 3561
Qy 641 GluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 659
Db 3562 GAGCTCCTGAGAAATCGAGACATATATTTAAAGTTAGTGAATGCACAGTCAGTGCAG 3618
RESULT 5
AX339031 AX339031 3702 bp DNA linear PAT 09-JAN-2002
LOCUS Sequence 13 from Patent WO0194400.
DEFINITION AX339031
ACCESSION AX339031
VERSION AX339031.1 GI:18129123
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM
REFERENCE 1
AUTHORS Frank, M.H. and Sayegh, M.H.
TITLE A gene encoding a multidrug resistance human p-glycoprotein
homologue on chromosome 7p15-21 and uses thereof
JOURNAL Patent: WO 0194400-A 13 13-DEC-2001;
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES
source Location/Qualifiers
1..3702
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 1098 a 708 c 856 g 1039 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 3,92e-259 Length: 3702
Score: 3323.00 Matches: 659
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Db: 6 Gaps: 0
US-09-873-409-1 (1-659) x AX339031 (1-3702)
Qy 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrSer 20
Db 1723 ATGCTGGCGGAGAAAGGAGCACATGCTGAACTAATGGCAACGAGGTCTATATTATTC 1782
Qy 21 LeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSer 40
Db 1783 CTTGTGATCTCAGAGATATTAAAAAGCTGATGAACAGATGGAGTCAATGACATATTCT 1842
Qy 41 ThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPhe 60
Db 1843 ACTGAAAGAAAGAACCAACTCACTTCTCTGCACTCTGTGAAGACATCAAGTCAGACTTC 1902
Qy 61 IleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeu 80
Db 1903 ATTGACAGGCTGAGGATCCACCAATCTAAAGATGAAGTCTTCTCTGAGTCTCTCTA 1962
Qy 81 LeuLysIleLeuLysLeuAsnLysProGluTyrProPheValValLeuGlyThrLeuAla 100
Db 1963 TTAATAATTTTAAAGTTAAACAAGCTGAATGCCCTTTTGTGTTCTGGGACATTTGGCT 2022
Qy 101 SerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThr 120
Db 2023 TCTGTTTAAATGGAACTGTTTCATCCAGTATTTTCCATCATCTTTGCAAAAATTTATAACC 2082
Qy 121 MetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIle 140
Db 2083 ATGTTTGGAAATTAATGATAAAACCACTTAAAGCATGATGACAGAAATTTATTCATGATA 2142
Qy 141 PheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGly 160
Db 2143 TTCGTCAATTTGGGTGTTATTGCTTTGTCAGTTATTTTCATGCAGGATTTATTTACGGC 2202
Qy 161 ArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyr 180


```
Db 2203 AGAGCAGGGGAAATTTAAACGATGAGATTAAAGACACTTGGCGCTTCAAAAGCCATGTTATAT 2262
Qy 181 GlnAepIIeAlaTtrPheAspGluLysGluAenSerThrGlyGlyLeuThrThrIleLeu 200
Db 2263 CAGGATATTGCTGTGTTGATGAAAGGAAACACACACAGAGGGCTTGACAAATATTA 2322
Qy 201 AlalleAspIIeAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGln 220
Db 2323 GCCATAGATATAGCAAAATTCAGGAGCAACAGGTTCCAGGATTTGGCGTCTTTACACAA 2382
Qy 221 AenAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrrPheMetThr 240
Db 2383 AATGCAACTTAACATGGGACTTTTCAGTTATCATTTTCTCTTATATATGAGTGGAGATGACA 2442
Qy 241 PheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAla 260
Db 2443 TTCCTGATTCGAGTATTGCTCCAGTACTTGCCTGACACAGGATGATTGAAACCCGACGA 2502
Qy 261 MetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThr 280
Db 2503 ATGACTGGATTGGCCAAACAAGATTAAGCAAGAACTTAAGCATGCTGGAAGATAGCAACT 2562
Qy 281 GluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGln 300
Db 2563 GAAGCTTTGGAGAAATATACGTACTACTAGTGTCTATTAAACAAGGGGAAAAAGCCCTTCGAGCAA 2622
Qy 301 MetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIle 320
Db 2623 ATGTATGAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCACAGATTATT 2682
Qy 321 GlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArg 340
Db 2683 GGAAGCTGTTATGCATTTCAGCCATCCCTTTATATATTTTGGCTATGCAGCAGGGTTTCGA 2742
Qy 341 PheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPhe 360
Db 2743 TTTGGAGCCTATTAAATTCAGCTGGAGCAATGACCCAGAGGGCATGTTTCATAGTTTTT 2802
Qy 361 ThrAlaIleAlaTyrGlyValaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyr 380
Db 2803 ACTGCAATTCGATATGGAGCTATGGCCATCGGAAAGACCTCGTTTGGCTCCTGCAATAT 2862
Qy 381 SerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIle 400
Db 2863 TCCAAAGCCAAATCGGGGCTCGCATCTGTTTGGCTTGTGGAAGAAAGAAACCAATATA 2922
Qy 401 AspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArg 420
Db 2923 GACAGCCGCAGTCAAGAAGGGGAAAAAGCCAGACACATGTGAAGGGAAATTTAGAGTTTCGA 2982
Qy 421 GluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeu 440
Db 2983 GAAGTCTCTTTCTATCCATGTCGCCAGATGTTTTCACTCCGGGCTTATCCCTC 3042
Qy 441 SerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThr 460
Db 3043 AGTATTGAGCGAGGAAGACAGTAGCATTTTGTGGGAGCAGCGGTGTGGGAAAGCACT 3102
Qy 461 SerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGly 480
Db 3103 TCTGTTCACATTTCTCAGAGACTTTTACCCCGTGCAGAGCAAGTGTGTTGATGGT 3162
Qy 481 ValAspAlaLysGluLeuAsnValGlnTrrPheArgSerGlnIleAlaIleValProGln 500
Db 3163 GTGGATGCAAAAGAAATTAATGTACAGTGGCTCCGTTCCCAATAGCAATCGTCTCTCAA 3222
Qy 501 GluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArg 520
Db 3223 GAGCCTGTGCTCTTCAACTGACAGCATTGTGAGAACATCGCTTATGTTGTCACACAGCGT 3282
Qy 521 ValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIle 540
```

```
Db 3283 GTGGTGCCATTAGATTGAGATCAAAAGAGCCCAAAATGCAGCAAAATATCCATTCTTTATT 3342
Qy 541 GluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGly 560
Db 3343 GAAGGTCTCCCTGAGAAATACACACACAAGTTGGACTGAAAGGAGACACAGCTTTCTGGC 3402
Qy 561 GlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeu 580
Db 3403 GGCCAGAAACAAAGACTAGCTATTGCAAGGGCTCTTCTCCAAAAACCCCAAAATTTTATTG 3462
Qy 581 LeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeu 600
Db 3463 TTGATGAGGCCACTTCAGCCCTCGATATACAGTGAAGAGTGGTTCAGATGCCCTT 3522
Qy 601 AspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGln 620
Db 3523 GATAAGCCAGCAGCGGAAGGACATGCTAGTGTCTCACTCACAGCTCTCTGCAATTCCAG 3582
Qy 621 AsnAlaAspLeuIleValValLeuHisGlnGlyLysIleLysGluGlnGlyThrHisGln 640
Db 3583 AACGCAGATTGTAGTGGTTCTGCACAATGGAAGATAAAAGGAACCAAGGAACCTCATCA 3642
Qy 641 GluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 659
Db 3643 GAGCTCTCGAATAATCGAGACATATATTTTAAGTTAGTGAATGCACAGTCACTAGTCAG 3699

RESULT 6
AX478104 3699 bp DNA linear PAT 12-AUG-2002
LOCUS
DEFINITION Sequence 34 from Patent WO0240541.
ACCESSION AX478104
VERSION AX478104.1 GI:22217064
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1
AUTHORS
Tang, Y.T., Yue, H., Nguyen, D.B., Hafalia, A.J., Elliott, V.S., Lu, Y.,
Walia, N.K., Yao, M.G., Baughn, M.R., Gandhi, A.R., Ding, L.,
Sanjanwala, M., Ramkumar, J., Arvizu C., Gietzen, K.J., Lal, P.G.,
Azimzai, F., Khan, F.A., Thangavelu, K., Thornton, M., Lu, D.A.,
Tribouley, C.M., Warren, B.A., Ison, C.H., Das, D., Raumann, B.E.,
Pollicy, J.L. and Kearney, L.
TITLE
Transporters and ion channels
JOURNAL
Patent: WO 0240541-A 34 23-MAY-2002;
Incyte Genomics, Inc. (US)
FEATURES
source
1. .3699
/organism="Homo sapiens"
/db xref="taxon:9606"
/note="Incyte ID No: 7472030CB1"
BASE COUNT 1116 a 707 c 860 g 1016 t
ORIGIN
```

```
Alignment Scores:
Pred. No.: 3,44e-246 Length: 3699
Score: 3163.00 Matches: 632
Percent Similarity: 96.05% Conservative: 1
Best Local Similarity: 95.90% Mismatches: 0
Query Match: 95.19% Indels: 26
DB: 6 Gaps: 1
```

US-09-873-409-1 (1-659) x AX478104 (1-3699)

```
Qy 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
Db 1798 ATGTGGCGGAGAAAGGCGCATGCTCAACTAATGGCAAAACGAGGCTCTATATTATCA 1857
Qy 21 LeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSer 40
Db 1858 CTTGTGATGTACAGGATATTTAAAAAAGCTGATGAACAGATGGAGTCAATGACATATCTT 1917
```


Alignment Scores:

Pred. No.: 5,8e-146 Length: 3924
Score: 1924.50 Matches: 372
Percent Similarity: 74.10% Conserved: 123
Best Local Similarity: 55.69% Mismatches: 162
Query Match: 57.91% Indels: 11
DB: 6 Gaps: 3

US-09-873-409-1 (1-659) x AX024455 (1-3924)

Qy 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrSer 20
Db 1851 GTAAATGTTGGAGCAAGGAGCCACAGCACTGATGAAGAGAGGGGTGACTTCAAA 1910
Qy 21 LeuVal-----MetSerGlnAspIleLysLysAlaAsp 31
Db 1911 CTTGTCAACATGCACACATCAGGAAGCCAGATCCAGTCAGAGAATTTGAACATAATGAT 1970
Qy 32 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis 51
Db 1971 GAAAGGTCGCCACTAGAAATGCCCAAAATGGCTGGAATCTCGCTATTAGGCATTCT 2030
Qy 52 SerValLysSerIleLys---SerAspPheIleAspLysAla---GluGluSerThrGln 69
Db 2031 ACTCAGAAAAACCTTAAAAAATTCACAAATGTGTCAAGAGAGCCTTGATGTGGAACCCGAT 2090
Qy 70 SerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysPro 89
Db 2091 GGACTTGAAGCAAAATGTGCCACAGATGCTTCTTGAAGGTCTTGAACCTGAATGAATAACA 2150
Qy 90 GluTrpPropheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPro 109
Db 2151 GAATGGCCCTACTTGTCTGGGAACAGATGTGCCATTGCCAATGGGGGCTTCAGCG 2210
Qy 110 ValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAspLysThrThr 129
Db 2211 GCATTTTTCAGTCATATCTCAGAGATCATAGCGATTTTGGACACAGCGATGATGATG 2270
Qy 130 LeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPhe 149
Db 2271 AAGCAGCAGAAAGTGAACATATCTCTTTGATTTCTTATTTCTGGAAATATTTCTTT 2330
Qy 150 ValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArg 169
Db 2331 TTTACTTCTCTCTCAGGGTTTCAGTTTGGAAAGCTGGCGAGATCTCTCACCAAGA 2390
Qy 170 LeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLys 189
Db 2391 CTGCGTCAATGGCTTTAAAGCAATGCTAAGACAGGACATGAGTGTGTTGATGACCAT 2450
Qy 190 GluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGly 209
Db 2451 AAAAACAGTACTGGTGCACTTCTCAAGACTTGCACAGATGCTGCCCAAGTCCAAGA 2510
Qy 210 AlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerVal 229
Db 2511 GCCACAGGACAGGTGGCTTTAATGACAGAAATATAGCTAACTTGGAACTGGTATT 2570
Qy 230 IleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProVal 249
Db 2571 ATCATATCATTTATCTACGGTTGGCAGTTAAACCTATTGCTATTAGCAGTTGTTCCAA 2630
Qy 250 LeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLys 269
Db 2631 ATTGCTGTGTGAGAAATGTTGAAATGAAATTTGGCTGGGAATGCCCCAAAAGAGATAA 2690
Qy 270 GlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIle 289
Db 2691 AAAGAACTGGAACTGTGGAAGATTGCAACAGAGCAATAGAAAAATATTAGGACAGTT 2750
Qy 290 ValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGln 309
Db 3831 TTTTCAATGGTGTGTCAGGCT 3854

Db 2751 GTGTCTTTGACCCAGGAAAGAAATTTGAATCAATGATGTTGAAAAATTTGATGGACCT 2810
Qy 310 HisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAla 329
Db 2811 TACAGGAATTTCTGTGCAGAGGCACACATCTATGGAATTAATCTTTAGTATCTCACAA 2870
Qy 330 PheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGly 349
Db 2871 TTTATGATATTTTCTATGCGGTTGTTTTCGATTTGGTGCATATCTCATTTGTGAATGA 2930
Qy 350 ArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyValMetAla 369
Db 2931 CATATGCCCTTCAGAGATGTTATTTCTGTGTTTCTGCAATTTGATTTGGTGCAGTGCT 2990
Qy 370 IleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHis 389
Db 2991 CTAGGACATGCCAGTTCATTTGCTCCAGACTATGCTAAAGCTAAGCTGCTCTGCAGCCCAC 3050
Qy 390 LeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLys 409
Db 3051 TTATTCATGCTGTTTGAAGACAACCTCTGATTGACAGCTACAGTGAAGAGGGGCTGAAG 3110
Qy 410 ProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArg 429
Db 3111 CCTGATAAATTTGAAGGAAATATAACATTTAATGAAGTCGTGTTCAACTATCCACCACGA 3170
Qy 430 ProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAla 449
Db 3171 GCAAACTGTCAGTGTCTCAGGGCTGAGCCTGGAGGTGAAGAAAGGCCAGACACTAGCC 3230
Qy 450 PheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyr 469
Db 3231 CTGTGGGACAGTGGCTGTGGGAAGACGACGCGTGTGTCAGCTCTCGAGCGGTCTTAC 3290
Qy 470 AspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGln 489
Db 3291 GACCCCTTGGCGGAGACAGTCTCTCAGTGTCAAGCAAGCAAGAAACTCAATGTCCAG 3350
Qy 490 TrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIle 509
Db 3351 TGGCTCAGACTCAACTCGGAATCGTGTCTCAGGAGCCTATCCTATTATTGACTGCAGCAT 3410
Qy 510 AlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGlu 529
Db 3411 GCCGAGAATATTGCTATGAGACAAACAGCGGGTGTATCACAGGATCAAAATTTGTGAGT 3470
Qy 530 AlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThr 549
Db 3471 GCAGCAAGCTGCCAACATACATCTCTTCATCAGAGCTTACCCCAAAATATGAAACA 3530
Qy 550 GlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAla 569
Db 3531 AGAGTGGAGATATAGGGAGCTCAGCTCTCAGAGGTCAAAAACAGAGGATTTGCTATTGCC 3590
Qy 570 ArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAsp 589
Db 3591 CGAGCCCTCATCAGACAACCTCAAAATCTCTCTGTTGGATGAAGCTACATCAGCTCGAT 3650
Qy 590 AsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCys 609
Db 3651 ACTGAAGTGAAGGTTGTCCAAGACCCCTGGCAAAAGCCAGAGAAGGGCGCCACCTGC 3710
Qy 610 LeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHis 629
Db 3711 ATTGTGATTGTCACCGCTGTCCACCATCAGATGACAGATTANTAGTGTGTTTTCAG 3770
Qy 630 AsnGlyLysIleLysGluGlnGlyThrHisGlnLeuLeuArgAsnArgAspIleTyr 649
Db 3771 AATGGACAGTCAAGGACATGGCAGCATCAGCAGCTGCTGGCACAGAAAGGCATCTAT 3830
Qy 650 PheLysLeuValAsnAlaGlnSer 657
Db 3831 TTTTCAATGGTGTGTCAGGCT 3854


```

RESULT 8
AX409652
LOCUS AX409652 3924 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 2299 from Patent WO0229103.
ACCESSION AX409652
VERSION AX409652.1 GI:21442357
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE
AUTHORS Alvares, C., Horne, D., Peres-da-Silva, S. and Vockley, J. G.
TITLE Gene expression profiles in liver cancer
JOURNAL Patent: WO 0229103-A 2299 11-APR-2002;
GENE LOGIC INC (US)
FEATURES
source Location/Qualifiers
1..3924
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="EMBL/GenBank Accession No. M23234"
BASE COUNT 1145 a 790 c 977 g 1012 t
ORIGIN
Alignment Scores:
Pred. No.: 5.8e-146 Length: 3924
Score: 1924.50 Matches: 372
Percent Similarity: 74.10% Conservative: 123
Best Local Similarity: 55.69% Mismatches: 162
Query Match: 57.91% Indels: 11
DB: 6 Gaps: 3

US-09-873-409-1 (1-659) x AX409652 (1-3924)
QY 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrSer 20
DB 1851 GTAATTGTGGAGCAAGGAGCCACAGCACTGATGAAGAAGGAGGGGTACTTTCAA 1910
QY 21 LeuVal-----MetSerGlnAspIleLysLysAlaAsp 31
DB 1911 CTGTGCAACATGCAGACATCAGGAAGCCAGATCCAGTCAGAGAATTTGAATGAT 1970
QY 32 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis 51
DB 1971 GAAAGGCTGCCACTAGAAATGGCCCAATGGCTGGAATCTCGCCCTATTAGGCATTCT 2030
QY 52 SerValLysSerIleLys---SerAspPheIleAspLysAla---GluGluSerThrGln 69
DB 2031 ACTCAGAAAACCTTAAATAATCACAATAATGTGTCAGAAAGAGCCCTTGATGGAAACCGAT 2090
QY 70 SerLysGluIleSerLeuProGluValSerLeuLysIleLeuLysLeuAsnLysPro 89
DB 2091 GGACTTGAACCAATGTGCCACCACTGCTCTTCTGAGAGTCTGAACTGAATAAACA 2150
QY 90 GluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPro 109
DB 2151 GAATGGCCCTACTTTGTGCGGGAACAGTATGTGCCAATGGCGGCTTCAGCCG 2210
QY 110 ValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThr 129
DB 2211 GCATTTTCATGTCATATCTCAGAGATCATAGCGATTTTGGACCGCGCATGATGAGTG 2270
QY 130 LeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPhe 149
DB 2271 AACGACGAGAGTGCACATATCTCTTTGATTTTCTTATTTCTGGGAATATTTCTTTT 2330
QY 150 ValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArg 169
DB 2331 TTTTACTTTCTCTCTCAGGGTTTCACGTTTGGGAAAGCTGGCGGAGATCTCCACCAAGA 2390
QY 170 LeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLys 189

```

```

DB 2391 CTGCGGTCAATGGCTTTTAAAGCAATGCTAAGACAGGACATAGCTGGTTGATGACCAT 2450
QY 190 GluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGly 209
DB 2451 AAAACAGTACTGGTGCATTTCTACAGACTTGGCCACAGATGCTGCCAAGTCCCAAGA 2510
QY 210 AlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerVal 229
DB 2511 GCCACAGAAACCCAGGTGGCTTTAATTCACAGAAATATAGCTTAACCTTGGAACTGTATT 2570
QY 230 IleIleSerPheIleTyrGlyTrpGluMetThrPheIleLeuLeuSerIleAlaProVal 249
DB 2571 ATCATATCATTTATCTACGGTTGGCAGTTAACCTTATGCTATAGCAGTTGTTCAAT 2630
QY 250 LeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLys 269
DB 2631 ATTGCTGTGCAGGAATCTTGAATGAAATGTTGGTGGTGAATGTCACAAAGAGATAAA 2690
QY 270 GlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIle 289
DB 2691 AAAGAACTGGAAAGCTGCTGGAAGATTCACACAGAGCAATAGAAAATATTAGGACAGTT 2750
QY 290 ValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGln 309
DB 2751 GTGCTTTTGACCCAGAAAGAAATTTGAATCAATGATGTTGAAATTTGATGACCT 2810
QY 310 HisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAla 329
DB 2811 TACAGGAATCTGTGCAGAGGACACATCTATGCAATTTACTTTTAGTATCTCACAGCA 2870
QY 330 PheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGly 349
DB 2871 TTTATGTTATTTTCCATGTCGCGTTGTTTTCGATTTGGTGCATATCTCATTTGTAATGA 2930
QY 350 ArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrClyAlaMetAla 369
DB 2931 CATATGCGCTTCAGAGATGTTATTTCTGGTGTGTTCTCGAAATTTGATTTGGTCAGTGGCT 2990
QY 370 IleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHis 389
DB 2991 CTAGGACATGCCAGTTTCTTGTCCAGACTATGCTAAAGCTTAAGCTGTCTCAGCCCCAC 3050
QY 390 LeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLys 409
DB 3051 TTATTCATGCTCTTTTGAAGAACAACCTCTGATTGACAGTACAGTCAAGAGGGGCTGAAG 3110
QY 410 ProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArg 429
DB 3111 CCTGATAAATTTGAAGGAATAATAACATTTAATGAAGTCGTGTTCACTATCCACCCCA 3170
QY 430 ProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAla 449
DB 3171 GCAACGTCGACAGTCTTTCAGGGCTGAGCTGGAGGTGAAGAAAGGCCAGACATAGCC 3230
QY 450 PheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyr 469
DB 3231 CTGTGGGCGAGCAGTGGCTGTGGGAAGACAGCGTGGTCCAGCTCTCTGGAGCGGTCTTAC 3290
QY 470 AspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGln 489
DB 3291 GACCCCTTGGCGGGGACAGTGGCTTCTCATGTGTCAAGAAAGCAAGAAATCAATGTCCAG 3350
QY 490 TrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIle 509
DB 3351 TGGCTCAGAGCTCACTCGGAATCGTGTCTCAGGAGCTATCTTATTTTACCTGACGATT 3410
QY 510 AlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGlu 529
DB 3411 GCCGAGAATATTGCTATGGAGACAACACCGCGGTGTTATCACAGGATGAATTTGTGAGT 3470
QY 530 AlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThr 549
DB 3471 GCAGCCAAAGCTGCCAACATACATCTTTTCATCGAGAGCTTACCCCAACAATATGAAACA 3530

```



```
Qy 550 GlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaLeu 569
:::|||||
Db 3531 AGAGTGGAGAGTAAGGGGACTCAGCTCTCAGGAGGTCAAAAACAGAGGATTCCTATTGCC 3590
Qy 570 ArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 589
Db 3591 CGAGGCCCTCATCAGCAACCTCAATCTCCCTGTTGGATGGAAGCTACATCAGCTCTGGAT 3650
Qy 590 AsnArgSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCys 609
Db 3651 ACTGAAAGTGAAGAGTTGTCAGAGNAGCCCTGGCAAGAGCCAGAGAGCCGACCTGTC 3710
Qy 610 LeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHis 629
Db 3711 ATTGTGATTGCTCACCGCTGTCACCATCCAGATGCAGACTTAAATAGTGGTGTTCAG 3770
Qy 630 ArgGlyLysIleLysGlnGlnGlyThrHisGlnGlnLeuLeuLeuLeuLeuLeuLeuLeu 649
Db 3771 AATGGAGAGTCAAGGAGCATGGCAGCATGCAGCATGCTGCTGGCAGAAAGGCATCTAT 3830
Qy 650 PheLysLeuValAsnAlaGlnSer 657
Db 3831 TTTTCATGCTCAGTCTCAGGCT 3854

RESULT 9
LOCUS HUMMDR3 3924 bp mRNA linear PRI 11-JUN-1993
DEFINITION Human membrane glycoprotein P (mdr3) mRNA, complete cds.
ACCESSION M23234
VERSION M23234.1 GI:187501
KEYWORDS P-glycoprotein; membrane glycoprotein.
SOURCE Human liver, cDNA to mRNA, clone 3.27.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 3924)
AUTHORS van der Blik,A.M., Koolman,P.M., Schneider,C. and Borst,P.
TITLE Sequence of mdr3 cDNA encoding a human P-glycoprotein
JOURNAL Gene 71 (2), 401-411 (1988)
MEDLINE 89138016
PUBMED 2906314
COMMENT Draft entry and computer-readable sequence for [1] kindly provided
by P. Borst, 21-MAR-1989.
FEATURES
Location/Qualifiers
1..3924
/organism="Homo sapiens"
/db_xref="taxon:9606"
<1..3924
/product="mdr3 mRNA"
33..3872
/note="P-glycoprotein"
/codon_start=1
/db_protein_id="AAA3207.1"
/db_xref="GI:307181"
/translation="MDLEAAKNGTAWRPTSAEGDFELGTSKQKREKTKTKVMIGVLT
LPRYSDWDLKPLMSLGTIMATAGSGLPLMMIVFGEMTDKFDVDTAGNSSFVPNFSLSL
LNPGLLEEMTRYAVYVYSGLGAVLVAAYIQVSWFTLAAGROIRKIRKQKPFHALRQ
EIGWFDINDTLENLRLDDISKISEGIDGKVMFFQAVAFAGFVGFIRGWLTL
VIMASPLIGLSAAWAKILSAFSDKLAAYAKAGAAVEALGAIKRTVIAFGQNKEL
ERYQHLNKAKEIGIKKAISSANISMGIAFLIYASALAFWYGTSLVLSKEYTTIGNAM
TVFFSILIGAFSGQAAVPCIDAFANARGAAVVIDIIDNNPKIDSFSRGRHKPDSIKG
NLEFNDVHFSYPSRANVKILGLNLKQSGVTVALVSGGCKSTTVTLVLIORLYDPDE
GTINIDGDIRNFNVNLYLREILGVVSQBPVLEFSTTIAENICYGRGNVTMDKIKAVKE
ANAYEFIMKLPKQFDTLVGERGAQSQGGOKRIARALVRNPKILLDEATLSALDTE
SNAEVOALDKPDKREGRTTIVIAHRLSTVRNADVIAGPDEGVIVEGSHSELMKKEGVY
FKLVNMQTSGSGIQSEEFELNDEKATMAPNGWKSRLFRSTQNLKNQSQMCKSLD
VETDLEANVPVPSFLVKLNLKNTWPYFVVGTCVCAIANGQLQAPAFVIFSEIIAIFG
PGSDAVKQKQKNI FSLIFLFLGIISFTFFLQGFQKAGEILTRLRSMFAFKMLRQ
DMSWFDHNSGALSTRLADAAQVQCATGTGRLALIAQNTANLGTGIIISIFWQL
TLLLIADVPIITAVSGIVEMKLAGNAKDEAKIATIAENIRNTVSLTQERK
FESVMYKLYGIPRNSVQKHIYGTITSIQAFMYFSYAGCFRFGAYLIVNIGHMFRD
VILVPSAIVFGAVALGHASSFAPYAKAKLSAAHLFMLFERQPLIDSYSEELGPKDF
```

```
EGNITFNEVVPYPTPRANVPVLOGLSLEVKKGOTLALVSSGCGKSTVQLBERPYPDP
LAGTVLLDGGAKKLNVQWLRQAQIGIVSQBFIPLFDCSIAENIATIGDMSRVVSQDEIYS
AAKAAHLHPFETLPHKYETRVGDKQRIATARALRQIPOILLDDATSA
LDTSEKVSQVQALDKAREGRTICVIAHRLSTIQNADLIVFQNGRVKEHGTQOULLAQ
KGIYFSKVSQVQAGTONL"
BASE COUNT 1145 a 790 c 977 g 1012 t
ORIGIN
Alignment Scores:
Pred. No.: 5,8e-146 Length: 3924
Score: 1924.50 Matches: 372
Percent Similarity: 74.10% Conservative: 123
Best Local Similarity: 55.69% Mismatches: 162
Query Match: 57.91% Indels: 11
DB: 9 Gaps: 3
US-09-873-409-1 (1-659) x HUMMDR3 (1-3924)
Qy 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
Db 1851 GTAAATTGTGGAGCAAGGAAGCCACAGCAACTGATGAAGAAGGAAGGGGTGACTTTCAAA 1910
Qy 21 LeuVal-----MetSerGlnAspIleLysLysAlaAsp 31
Db 1911 CTGTGCAACATGCAGACATCAGGAAGCCAGATCCAGTCAGAAGAAATTTGAACATAATGAT 1970
Qy 32 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis 51
Db 1971 GAAAAGGCTGCCACTAGAAATGGCCCAAAATGGCTGAAATCTCGCTCTATTAGGCATTCT 2030
Qy 52 SerValLysSerIleLys---SerAspPheIleAspLysAla---GluGluSerThrGln 69
Db 2031 ACTCAGAAAAACCTTAAAAAATTCACAAATGTGTCAAGAAGAGCCTTGATGTGGAAAACCGAT 2090
Qy 70 SerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysPro 89
Db 2091 GGACTTGAAGCAAAATGTGCCACACAGTGTCTTCTTGAAGGTCCTGAAACTGAATATAACA 2150
Qy 90 GluTrpProPheValIleuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPro 109
Db 2151 GAATGGCCCTACTTTGTCTGTGGAAACAGTATGTGCCATGGCCATGGGGGGCTTCAGCGG 2210
Qy 110 ValPheSerIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThr 129
Db 2211 GCATTTTTCAGTCATATCTCTCAGAGATCATAGCGATTTTGGACACAGGCGATGATGACGTG 2270
Qy 130 LeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPhe 149
Db 2271 AAGCAGCAGAGAGTGCAACATATCTCTTTGATTTTCTTATTTCTGGAAATATTTCTTTT 2330
Qy 150 ValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArg 169
Db 2331 TTTACTTTCTCTTCAGGGTTTCAGTTTGGGAAAGCTGGCGAGATCTCTACCAGAGA 2390
Qy 170 LeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLys 189
Db 2391 CTCGCGTCAATGGCTTTTAAAGCAATGTCTAAGACAGGACATGAGCTGTGTTGATGACCAT 2450
Qy 190 GluAsnSerThrGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGly 209
Db 2451 AAAAAACAGTACTGGTGCACTTTCTACAAGACTTGCACAGATGCTGCCCAAGTCCAAGA 2510
Qy 210 AlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerVal 229
Db 2511 GCCACAGNACAGGTGGCTTTAATTCACACAGATATAGCTTAACCTTGGAACTGGTATT 2570
Qy 230 IleIleSerPheIleTyrGlyTrpGluMetThrPheLeuLeuLeuSerIleAlaProVal 249
Db 2571 ATCATATCATTTATCTACGGTTGGCAGTTAACCCCTATTGCTATTAGCAGTTGTTCCAATT 2630
Qy 250 LeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLys 269
Db 2631 ATTGCTGTGTCAGGAATTGTTGAAATGAAATTTGTTGGCTGGAAATGCCCAAAAGAGATAAA 2690
```


us-09-873-409-1 (1-659) x CRUPGP1185 (1-3987)

```
Qy 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1597 GTCAATTTGGGCAAGAAATCATCAAGAGCTCATCAAGAGAGGAGGCAATTTACTTCAA 1656
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 21 LeuValMetSerGln-----AspileLysAla 30
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1657 CTTGTCATGACACACACAGCAGCAAGAAATGAAATTTGAATTAGGAAATGAAAGTTGGTGAGTCT 1716
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 31 AspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeu----- 48
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1717 AAAAATGAAATTTGATTAATTTAGCATGCTCTCAAGAGATTACAGATCAGCTCAATATCAGA 1776
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 49 ProLeuHisSerValLysSerLysSerAspPheIleAspLysAlaGluGluSerThr 68
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1777 AGGAGATCAACTCGCAGAGATATCCGTGGACACATGACCAAGACAGCAAGCTTAGTACC 1836
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 69 ---GlnSerLysGluLysSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 87
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1837 AAAGAGGCGCTTGGATGAAGATGTACTCCAAATTTCTTTTGGCGTATCCTAAAGTTGAAT 1896
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 88 LysProGluTyrProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1897 TCATCTGAATGCCATATTTTGTGTGTGTATATCTGTGCCATAGTAAATGGACCTTTG 1956
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 108 HisProValPheSerIleIlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAsp 126
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1957 CAACAGCATTTCTCAATAATTTTCTAAGGTTGTAGGGGTTTTCACAAGAAATACTGAT 2016
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 127 LysThrThrLeuLysHisAspAlaGluLysTyrSerMetIlePheValIleLeuGlyVal 146
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2017 GATGAACCAACACGATGATAGCAACTTGTCTTCTTATTTCTGATCCTTGGGGTC 2076
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 166
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2077 ATTTCTTTTATTACATTTTCTTCAGGGCTTCACATTTGGCAAGCTGGAGAGATCCTC 2136
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspileAlaIlePhe 186
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2137 ACCAAGCGACTCCGATACATGCTTTTCAATCCATCTGAGACAGGATGTGAGCTGGTTT 2196
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 187 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspileAlaGln 206
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2197 GATAACCTTAAAAACACCACTGGAGCATTTGACCAAGCTCGCCAAACGATCCTGGTCAA 2256
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 207 IleGlnGlyAlaThrClySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2257 GTTAAAGGGGCTACAGGAGCCAGACTTGTCTATACCAAGAAATATATAGCAAAATCTTGGG 2316
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 227 LeuSerValIleIleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIle 246
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2317 ACAGAAATCATATATCCCTAATCTATGCTGGCAGTTGACACTTCTACTCTTAGCAAT 2376
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2377 GTCCCATCATTTGCAATAGCAGGAGTGTGTAGATGAAATTTGTCTGCAAGCACTA 2436
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 267 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2437 AAAGATTAAGAGGAGCTAAGAGGTTCTGGGAAGATTGCTTACTGAAGCAATAGAACTTC 2496
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2497 CGCATGTCTCTCTTTGACTCGGGAGCAAGATTGAAATATATGATGCCAGAGCTTTG 2556
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2557 CAGATACCATACAGAAATGCTCTGAAGAAAGCACACGCTCTTTGGGATCACTTCTCCTTC 2616
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 346
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2617 ACCCAGGCCATGATGATTTTTTTCGATGTCTGTTTCCGGTTTGGTGTACTTACTTGGTG 2676
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
Qy 347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2677 GCACGGGAATTAATGACATTTGAAAAATGTTCTATTAGTATTCTCAGCTATTGTTCTTGGT 2736
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2737 GCCATGGCAGTTGGCGAGTTCAGTTCTTCTGACTATGCCAAAGCCAAAGTGTCTG 2796
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 387 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 406
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2797 GCATCGCATATCATCATGATCATTTGAAAAAGTCCCTTCCATTGACAGCTACAGCACCGGA 2856
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 407 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 426
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2857 GGCCTGAAGCCCTAATACATTTGGAAGAAATGTAATTAATGAAGTCTGTTCTCAACTAC 2916
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2917 CCCACCGGCTGACATCCCATGCTTCAGGGGCTGAACCTGGAGGTGAAGAAGGGCCAG 2976
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuGln 466
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2977 ACGTGGCCCTGTTGGCAGCAGTGGCTGCGGGAAGACACAGTGGTCCAGCTGCTCGAG 3036
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 467 ArgLeuTyrAspProValGlnGlyValLeuPheAspGlyValAspAlaLysGluLeu 486
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3037 CGGTCTACGACCCCATGGCCGGAACAGTGTCTAGATGGCAAGAGTAACCAACTA 3096
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 487 AsnValGlnTyrLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3097 AATGTCCAGTGGCTCCGAGCACACTAGGCATTTGTCTCCCAAGAGCCCATCTGTTTCA 3156
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValProLeuAspGlu 526
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3157 TGCAGCATCGCTGAGACATCGCTTATGGACACAAACAGCCGGTCTGTCGCCAGATGAG 3216
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 527 IleLysGluAlaAlaAsnAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3217 ATCGAGAGGGCGGCAAGAAAGCTTCTCCAGGAAGCTCTGGACAAAGCCAGAGAAGGC 3276
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 547 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 566
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3277 TACAACACCAAGTAGGAGACAAAAGGGACTCAGCTGTCGGTGGGCAAGAAACAAACGATC 3336
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuAspLysAlaArgThrGly 586
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3337 GCCATAGCACGGCCCTCGTCAGACAGCTCACATTTGCTTCTGGATGAAGCAACATCA 3396
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 587 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3397 GCCCTGGATACAGAGAGTGAAGGTTCTCCAGGAAGCTCTGGACAAAGCCAGAGAAGGC 3456
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspIleVal 626
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3457 GGCACCTGATTTGATGCTCCACGGCTGTCCACCATCCAGAACGACAGACTTGTATCGTG 3516
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 646
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3517 GTGATTCAGATGGCAGGTCAAGGAGTGGACCCACAGAGCTGCTGGCACAGAAA 3576
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 647 AspileTyrPheLysLeuValAsnAlaGlnSer 657
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3577 GGCATCTATTTCTCCATGCTCAGTGTGCGAGCT 3609
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 11

```
CRUPGP1
LOCUS 4296 bp mRNA linear ROD 22-JUN-1995
DEFINITION C.griseus P-glycoprotein (isoform class I) mRNA, complete cds.
ACCESSION M60040
VERSION M60040.1 GI:191164
KEYWORDS multidrug resistance; p-glycoprotein; transmembrane protein.
SOURCE C.griseus adult liver and ovary, cdna to mRNA.
```


Db	2086	AGAGATCAACTCGCAGAAAGTATTCGTGGACACCATGACCCACGACGAGCAAGCTTAGTACC	2145
Qy	69	---GlnSerIysGlnIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn	87
Db	2146	AAAGAGCGCTTGGATGAAGATGTAACCTCAATTCTTTTGGCGTATCCTAAAGTTGAAT	2205
Qy	88	LysProGluTriProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal	107
Db	2206	TCATCTGAATGGCCATATTTTGTGGTGTGTATATCTGTGCCATAGTAAATGGAGCCCTTG	2265
Qy	108	HisProValPheSerIleIlePheAlaLysIleIleThrMetPhe--GlyAsnAsnAsp	126
Db	2266	CAACCCAGCATCTCAATAATATTTCTAAGTGTAGGGGTTTTTCAAGAATACTGAT	2325
Qy	127	LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal	146
Db	2326	GATCAAAACCAACGATCATGACCACTGTGTTTCTTATTGTTCTGATCCTTGGGGTC	2385
Qy	147	IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu	166
Db	2386	ATTTCCTTTTATACATTTTTTCTTCAGGCGCTTCAATTTGGCAAGCTGGAGAGATCCTC	2445
Qy	167	ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe	186
Db	2446	ACCAAGGCATCCGATACATAGTTTTCAAATCCATGCTGACAGAGATGTGAGCTGGTTT	2505
Qy	187	AspGlnLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln	206
Db	2506	GATACCCCTAAAAACACCACTGGAGCATTTGACCACGAGCTGCCAACGATGCTGGTCAA	2565
Qy	207	IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly	226
Db	2566	GTTAAAGGGCTACAGCAGCCAGACTTGCTGTCTATCCCAATATAGCAAACTTTGGG	2625
Qy	227	LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle	246
Db	2626	ACAGGAATCATCATATCCCTAATCTATGCTGGCAGTTGACACTTCTACTCTTAGCAATT	2685
Qy	247	AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn	266
Db	2686	GTCCCATCATTCGATAGCAGAGTGGTTGAGATGAATAATGTTCTCTGGACAGCACTA	2745
Qy	267	LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle	286
Db	2746	AAAGATAAAGAGGAGCTAGAAAGTCTCGGAAGATGCTACTCTGAAGCAATAGAGAACTTC	2805
Qy	287	ArgThrIleValSerLeuThrArgGlnLysAlaPheGluGlnMetTyrGluGluMetLeu	306
Db	2806	CGCACTGTCTGCTCTTTGACTCGGAGCAGAAAGTTTGAATAATATGTATGCCCCAGAGCTTG	2865
Qy	307	GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe	326
Db	2866	CAGATACCATAACAGAAATGCTCTGAAGAAGCAGACAGCTTTTGGGATCACATCTCTCTTC	2925
Qy	327	SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrIleuIle	346
Db	2926	ACCCAGGCCATGATGATTTTTCGTATGCTGTTTCCGTTTCCGTTTGGTGTCTTACTTGGTG	2985
Qy	347	GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly	366
Db	2986	GCACGGGAATTAATGACATTTGAAATGTTCTATTAGTATTTCTCAGCTATTGTCTTGGT	3045
Qy	367	AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly	386
Db	3046	GCCATGGCAGTTGGCAGGTCAGTTTCATTTGCTCTCTGACTATGCCAAGCCAAAGTGTGC	3105
Qy	387	AlaAlaHisLeuPheAlaLeuLeuGlyLysLysProAsnIleAspSerArgSerGlnGlu	406
Db	3106	GCATCGCATATCATCATCATTCATGAAAGTCCCTTCATTTGACAGCTACAGCAGCCGGA	3165
Qy	407	GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr	426
Db	3166	GGCCTGAAGCCTTAATACATTTGAAGAAATGGAATTTAATGAAGTCGTCTCAACTAC	3225


```
Qy 427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446
|||
|||
Db 3256 CCCACCCGGCTGACATCCAGTCTGCTCAGGGCTGAACCTGGAGTGAAGAAGGCCAG 3285
|||
|||
Qy 447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuGln 466
|||||
|||||
Db 3286 ACCTGGCCCTGGTGGGACGAGTGGCTGCGGGAAGAGACACAGTGGTCCAGCTGCTCGAG 3345
|||||
|||||
Qy 467 ArgLeuTyrAspProValGlnGlnValLeuPheAspGlyValAspAlaLysGluLeu 486
|||
|||
Db 3346 CGGTCTACGACCCATGGCCGAACAGTGTCTAGATGGCAAGAAGTAAACCACTA 3405
|||||
|||||
Qy 487 AenValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGlnProValLeuPheAen 506
|||||
|||||
Db 3406 AATGTCCAGTGGCTCCGAGCACACCTAGGCATTGTGTCCCAAGAGCCCATCTGTTTAC 3465
|||||
|||||
Qy 507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAenSerArgValValProLeuAspGlu 526
|||||
|||||
Db 3466 TGCAGCATCGCTGAGAACATCGCTTATGGAGAACACACCGCGGTGCTGTCCAGGATGAG 3525
|||||
|||||
Qy 527 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546
|||||
|||||
Db 3526 ATCGAGAGGGCGGCGAAGAACCAATCCACAGTTCATTGATCTACTGCTGCACAAA 3585
|||||
|||||
Qy 547 TyrAenThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 566
|||||
|||||
Db 3586 TACACACACAGAGTAGGACACAAAGGACTCAGCTGCTCGGTGGCGGACAAACACGCATC 3645
|||||
|||||
Qy 567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 586
|||||
|||||
Db 3646 GCCATAGCACGGCCCTCGTCAGACAGCCTCACATTTGCTCTCGATGAAGCAACATCA 3705
|||||
|||||
Qy 587 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606
|||||
|||||
Db 3706 GCCCTGGATACAGAGAGTGAAAGGTGTCCAGGAAGCTCTCGACAAAGCCAGAGAGGC 3765
|||||
|||||
Qy 607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626
|||||
|||||
Db 3766 CGCACTGCTATGTCATCGCTCACCGCTGTCACCATCCAGACGCGAGCTTGATCGTG 3825
|||||
|||||
Qy 627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlnLeuLeuArgAenArg 646
|||||
|||||
Db 3826 GTGATTCAGATGGCAAGTCAAGAGTGCACCCACAGCAGCTGCTGTCACAGAAA 3885
|||||
|||||
Qy 647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657
|||||
|||||
Db 3886 GGCATCTATTCTCCATGTCAGTGTGCAGGCT 3918
|||||
|||||
RESULT 12
CRUPGP1165 CRUPGP1165 4304 bp mRNA linear ROD 27-APR-1993
LOCUS Chinese hamster p-glycoprotein mRNA (clone ADX165), complete cds.
DEFINITION
ACCESSION M59253
VERSION M59253.1 GI:191154
KEYWORDS multidrug resistance glycoprotein; p-glycoprotein.
SOURCE Chinese hamster, cDNA to mRNA, clone ADX165.
ORGANISM Cricetulus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
Cricetulus.
REFERENCE 1 (bases 1 to 4304)
AUTHORS Devine,S.E., Huesain,A., Davide,J.P. and Melera,P.W.
TITLE Full length and alternatively spliced pgpl transcripts in
multidrug-resistant Chinese hamster lung cells
JOURNAL J. Biol. Chem. 266 (7), 4545-4555 (1991)
MEDLINE 91154265
PUBMED 1671863
FEATURES
Location/Qualifiers
source 1..4304
/organism="Cricetulus sp."
/db_xref="taxon:10031"
/cell_type="fibroblast"
```

```
/tissue_type="lung"
1..4304
/gene="pgp-1"
111..3941
/gene="pgp-1"
/codon_start=1
/product="p-glycoprotein"
/protein_id="AAA37004.1"
/db_xref="GI:191155"
/translation="MEPEEDFSGRKDKNFKLMGRKSKKEKKEKPPVSVFTMFRYAGW
LDRLMLVGTAAALIHGVALPLMLVFGDMTDFASVGNITPTNATNATNOASDIFG
KLEEMTYAYVYGTIGAGVLIVAYIOVSFMCCLAAGROJHKIRKQFPHAINQOIGWF
DVHDVGLNRLTDDVSKINEGIDKIMFQMATFFGGFICGTGKWLTLVLAI
SPVLGSAGIWAIXLISFTDKELQAYAKAGAAVEELAAITVIAFGGQKLEBRYN
NLEAEKRLGIKKAITANISMAAPLIIYASALAFWYGTSLVISKEYSIGVLTFFFA
VLTAPSIGQASNIEAFANARGAAEYFNIIDNKPISDFSKNGYKPDNIKLEKFL
NIHFSPSRKDVOILKGLNLKVGOSQTVLVNMSGCGKSTTVQLQRLYDPTGEVSI
DGODIRINVRILREIIGVVSQBPVLPAITIAENIRYGRNVTMDEIEKAKEANAYD
FIMKLPHKPTLHGERGAQLSGGQKORIAARALVRNPKILLDLDEATSALDTESEAVV
QAALDKAREGRTTIVIAHRLSTVRNADIAGFDGGVIVEQGNHEELMREKGIYFKLVN
TQAGNEIELNGESKNEIDNMSKDSASLSIRRRTRRSIRGPHDQDKLSKT
EALDEDVPTISFWRLILKNSSEWPYFVVGIFCAIVNGALQAPAFSIIIFSKVVGFTRTN
DDETKRHSNLSLPLILGIVISFTFFLOGFTFGKAGEILTTRKLRYMVFKSMURQDV
SWFDPNKYTGALTTRLANDAGQVKGATGARLAVITONIANLGTGIIISLYGWLTL
LLLAIVPIIAGVVMKLSGQALKDKKELEGSGKIAETAEINFRVTSLTRQKFE
NMYAQSLQIPYRNALKKARVFGTTSFTQAMYPFSYAAACFRFGAYLVARELMTFENVL
LVFSALIVFGAMVQSVSSFPADYAKAKVSASHIIMIIEKVPISDSYTGGLKPNLEGL
NVAFNEVFNVPTRPDIPLVLQNLNLEVKQGTALVAGSSGCGKSTVQLLERFYDPMA
GTFLDCKEYNQNLNVRALHGLIVSOEPIILFDCSIAENIAYGNSGVSQDEITERAA
KEANIHOFIESLDPKYNTRVGDGTQSGGQKORIAARALVRPHILLDLDEATSALD
TESEKVVQEAALDKAREGRTCTVIAHRLSTIQNADLIIVIQNGKVEHGTGQLLAQKG
IYFSMVSVQAGAKR"
BASE COUNT 1275 a 895 c 1031 g 1103 t
ORIGIN Map position lq26.
Alignment Scores:
Pred. No.: 3,19e-145 Length: 4304
Score: 1916.00 Matches: 360
Percent Similarity: 74.96% Conservative: 143
Best Local Similarity: 53.65% Mismatches: 154
Query Match: 57.66% Indels: 14
Db: 10 Gaps: 4
US-09-873-409-1 (1-659) x CRUPGP1165 (1-4304)
Qy 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrSer 20
|||||
|||||
Db 1914 GTCATTTGGAGCAGAGAAATCATGAAGGCTCATGAGAGAGGCGCATTTACTTCAA 1973
|||||
|||||
Qy 21 LeuValMetSerGln-----AspIleLysLysAla 30
|||||
|||||
Db 1974 CTTGTGATGACACAGACAGCAGGAAATGAAATTTAGGAATGAAGTTGGTGAGTCT 2033
|||||
|||||
Qy 31 AspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeu----- 48
|||||
|||||
Db 2034 AAAAATGAATTTAGATATTAGACATGTCTTCAAAAAGATTTCAGCATCCAGCTAATCAGA 2093
|||||
|||||
Qy 49 ProLeuHisSerValLysSerLysSerAspPheIleAspLysAlaGluGluSerThr 68
|||||
|||||
Db 2094 AGGAGATCAACTCGCAGAAAGTATCGTGGACCAATGACCAAGACAGGAAGCTTAGTACC 2153
|||||
|||||
Qy 69 ---GlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 87
|||||
|||||
Db 2154 AAAGAGGCTTGGATGAAGATGACTCCATTTCTTTGGCGTATCTCTAAAGTTGAT 2213
|||||
|||||
Qy 88 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107
|||||
|||||
Db 2214 TCATCTGAATGGCCATATTTTGTGGTGTGATATTTCTGTGGCATAGTAAATGCAGCTTG 2273
|||||
|||||
Qy 108 HisProValPheSerIlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAsp 126
|||||
|||||
Db 2274 CAACACGATCTTCAATAATATTTTCTAAGGTTGTAGGGTTTTTCAAGAAATACTGAT 2333
|||||
|||||
```


QY 127 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 146
 DB 2334 GATGAACCAACAGCATGATCACTGTTCTTCTTATTTCTGATCCTTTGGGTC 2393
 QY 147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 166
 DB 2394 ATTCTTTTATACATTTTCTTCAGGGCTTCACATTTGGCAAGCTGGAGATCCTC 2453
 QY 167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 186
 DB 2454 ACCAAGCGACTCCGATACATGTTTCAAATCCATGCTGAGACAGGATGTGAGCTGCTT 2513
 QY 187 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206
 DB 2514 GATAACCTTAAACACCACTGGAGCATTTGACCACGAGCTCGCCACCATGCTGTCAA 2573
 QY 207 IleGlnGlyValThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226
 DB 2574 GTTAAAGGGCTACAGAGCCAGACTTGTGTCATTTACCAGATATAGCAATCTTGGG 2633
 QY 227 LeuSerValIleIleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIle 246
 DB 2634 ACAGGAATCATCATATCCCTTAATCTATGCTGGCAGTTGACACTTCTACTCTTAGCAAT 2693
 QY 247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266
 DB 2694 GTCCCCATCATTCATGACAGGAGTGGTTGAGATGAAATGTTGTCGACAAGCACTA 2753
 QY 267 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286
 DB 2754 AAAGATAAGAGGAGCTAGAGGTTCTGGAGATTTGCTTACTGAACATATAGAAACTTC 2813
 QY 287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306
 DB 2814 CGCACTGCTCTCTTTGACTCGGAGCAGAGTGTGAAATATGATGCCAGACTTG 2873
 QY 307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326
 DB 2874 CAGATACCATACAGAATGCTCTGAAGAAAGCACACGCTCTTTGGGATCATCTCTCTTC 2933
 QY 327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 346
 DB 2934 ACCAGGCCATGATGATTTTTCGTATGCTGCTTGTTCGGTTGGTCTTACTTGGTG 2993
 QY 347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366
 DB 2994 GCACGGGAATTAATGACATTTGAAATGTTCTATTAGTATCTCAGCTATTGCTTGTGT 3053
 QY 367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386
 DB 3054 GCATGGCAGTTGGGAGGTCACTTCATTTGCTCTGACTATGCCAAAGCCAAAGTGTG 3113
 QY 387 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 406
 DB 3114 GCATCCCATATCATGATGATCATTTGAAAGTCCCTTCCATTGACAGCTACAGCCGGA 3173
 QY 407 GlyLysLysProAspThrCysGluGlyAsnLeuLeuGluPheArgGluValSerPheTyr 426
 DB 3174 GGCTGAAGCCTTAATACATTTGGAAGGAATGTGAATTTAATGAAGTCGTTCACATAC 3233
 QY 427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446
 DB 3234 CCCACCCGGCTGCATCCAGCTTCAGGGCTTCAAGCTGAGGTGAAGAGGGCCAG 3293
 QY 447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuGln 466
 DB 3294 ACCTGGCCCTGTGGGACAGCTGCTCGGGAAGAGCACAGTGTCTCAGCTGCTCGAG 3353
 QY 467 ArgLeuTyrAspProValGlnGlnValLeuPheAspGlyValAspAlaLysGluLeu 486
 DB 3354 CGGTTCTACAGCCCATGCCGGAACAGTGTCTTAGATGGCAAGAAAGTAAACCACTA 3413
 QY 487 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506

DB 3414 AATGTCCTGCTCGAGCACACCTAGGCATGTGTCCTCAAGAGCCCATCTGTTTGAC 3473
 QY 507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 526
 DB 3474 TCGAGCATGCTGAGACATCGCTTATGGAGACACAGCCGGTGTGTCCTCCAGATGAG 3533
 QY 527 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546
 DB 3534 ATCGAGGGCGCGCAAGGAAGCAACATCCACAGTTTCATTGAGTCACTGCTGCACAAA 3593
 QY 547 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlnLysGlnLeu 566
 DB 3594 TACAACACAGAGTAGGAGACAAAGGACTCAGCTGTCTGGTGGGAGCAAAACAGCATC 3653
 QY 567 AlaIleAlaThrAlaLeuLeuGlnLysProLysIleLeuLeuAspGluAlaThrSer 586
 DB 3654 GCCATAGCACGGCCCTCGTCAGACAGCTCACAATTTGCTCTGGATGAGCAACATCA 3713
 QY 587 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606
 DB 3714 GCCTTGGATACAGAGTGAAAGGTTGTCCAGGAAGCTCTCGCAAGCCAGAGAGGC 3773
 QY 607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626
 DB 3774 CGCACTGCATGTGATCGCTCACCGCTGTCCACCATCCAGACGCAAGACTTGATCGTG 3833
 QY 627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 646
 DB 3834 GTGATTCAGATGCGCAGGTCAAGAGGATGGCACCCACAGCAGCTGCTGCGACAGAAA 3893
 QY 647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657
 DB 3894 GGCACTTATTTCTCCATGCTCAGTGTGCGAGGT 3926
 RESULT 13
 MUSMDRA
 LOCUS
 DEFINITION Mouse mdr gene encoding a multidrug resistance protein mRNA,
 complete cds.
 ACCESSION J03398
 VERSION J03398.1 GI:1199109
 KEYWORDS multidrug resistance protein.
 SOURCE Mouse (strain BALB/c) pre-B-cell (cell line 702/3), cDNA to mRNA,
 clone MDR2.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (Bases 1 to 4084)
 AUTHORS Gros, P., Raymond, M., Bell, J. and Housman, D.
 TITLE Cloning and characterization of a second member of the mouse mdr
 gene family
 JOURNAL Mol. Cell. Biol. 8 (7), 2770-2778 (1988)
 MEDLINE 88302195
 PUBMED 3405218
 REFERENCE 2 (sites)
 AUTHORS Raymond, M. and Gros, P.
 TITLE Mammalian multidrug-resistance gene: correlation of exon
 organization with structural domains and duplication of an
 ancestral gene
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 86 (17), 6488-6492 (1989)
 MEDLINE 89367274
 PUBMED 2570420
 REFERENCE 3 (sites)
 AUTHORS Raymond, M. and Gros, P.
 TITLE Cell-specific activity of cis-acting regulatory elements in the
 promoter of the mouse multidrug resistance gene mdr1
 JOURNAL Mol. Cell. Biol. 10 (11), 6036-6040 (1990)
 MEDLINE 91042535
 PUBMED 2248681
 COMMENT Submitted in computer readable form by P.Gros 08-JUL-1988.
 FEATURES Location/Qualifiers
 source 1..4084

/organism="Mus musculus"
 /strain="BALB/c"
 /sub_species="domesticus"
 /db_xref="taxon:10090"
 /cell_line="702/3"
 /cell_type="pre-B cell"
 194..4024

CDS

/codon_start=1
 /product="multidrug resistance protein"

/protein_id="AAA39516.1"

/db_xref="GI:387428"

/translation="MDLEARNGTARLDGDFELGSLISNOGRKKKKVNLIGLTLFR
 YSDQKLFMLFTLMAIAHSGGLPLMMIVFGEMTDKFDVNTGNFSLVNFSLMLNP
 GRILEEMTYAYYSGLGGVLAAYIQVSPWTLAAGQIKKIRKQKFFHQLQWLMQ
 WFDIKGTTELRLTDDSVKISIGIGDKVGMFFQAIATPAGFIVGFIKWKLTIVIM
 AISPIGLSTAVAKILSTSDKELAAAYKAGAVEAPCAIRTVIAFGGKNELERY
 QKLELNAGKIGIKKISANISMGIAPELLIYASALAPFWGSLTVISKEVTIGNMTVF
 FSILIGAFVQQAACIDAFANARGAAVIFDIIDNNPKIDSFSEGRKPDNFKGLE
 FSDVHFSYPRANIKILGLNLKVKSGQTVALLVNGSCGKSTTVQLLRUDYDPEGKI
 SIDQDIRNFNVRCLREIIGVSOEPLVFTTIAENIRYGRNVMTDEIEKAVEANA
 YDFIMKLPQKDFLVDGRGAQLSGQKQRIAIARALVRNPKILLDEATSDLTSEEA
 EVOAALDKAREGRTTIVIAHRLSTIRNADVIAGFEDGVI VEOGSHSLMKKEGIYFRL
 VNMOTAGSOILSEFEVELSDEKAAGDVA PNGWKARFRNSTKSLKSPHONRLDET
 NELDANVPVPSFLVKLKLNTPEYFVGVTCALANGALOPAFSIIILSEMIALFGPD
 DAVQKCNMPSFLVLGVLSEFFFLQGTFGKAGEILTLRLRSWAFKMLRQDMS
 WFDHKNSTGALSTRLATDAQVQATGKTALIAQNTANLGTGIIISFYIQLTLL
 LLSVHPPIAVAGIVEMKAGNARKDKMEAKIATEAENIRTVSVISIQBRKTES
 MYVEKLVGRNVRSHAKHIIYGTISIOWAFYSYAGCFRFGSYLI VNGHMRPKDVL
 VFSNVLGVALVAFKAFADYAKAKLSAAYLPSLPERPLIDSYSGEGLPWPKFEGS
 VTNEVFNTPTRANVPVLOGLSLVKVKGQTLALVSSGCGKSTVQVLLERFDPMAG
 SVLLDQGEAKNQLWRQAGIVSQBPILPDCSIAENIAYGNSRVVPHDEIVRAAK
 EAMNHPPIETLPKYNTRVQDKTQLSGGQKRIARALIRQPRVLLLDSEATSDLT
 ESEKVOELDKAREGRTTIVIAHRLSTIQNADLIVVIENGKVEKHEHQQLLAQKI
 YFSMNIQAQTONL"

BASE COUNT 1117 a 934 c 1075 g 958 t
 ORIGIN 43 bp upstream of HinfI site.

Alignment Scores:

Pred. No.:	3,27e-145	Length:	4084
Score:	1915.50	Matches:	366
Percent Similarity:	73.95%	Conservative:	128
Best Local Similarity:	54.79%	Mismatches:	163
Query Match:	57.64%	Indels:	11
DB:	10	Gaps:	1

US-09-873-409-1 (11-659) x MUSMDRA (1-4084)

Qy	1	MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer	20
Db	2003	GTCAATTGTGGAACAAGGACGACAGTGTGATGAAGAAGGAGGATCTACTTCAGA	2062
Qy	21	LeuValMetSerGln-----AspIleLysLys	29
Db	2063	CTCGGTAAATGACACAGACAGGAGGACGATCTGTGACAGAAGATTGAAAGTTGAGCTA	2122
Qy	30	AlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuPro	49
Db	2123	AGTGACGAAGAGCTGCTGGAGATGTGCCCAATGGCTGGAAGACGACGATATTAGG	2182
Qy	50	LeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThrGln	69
Db	2183	AATTCTACAAAGAAAAGCTTAAAGTCCACATCAGATAGGCTGATGAAGAAACCAAT	2242
Qy	70	SerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysPro	89
Db	2243	GAACTTGATGCAACCGTCCACGATGTCTTTCTGAAGGTCTTAAAGTGAATAAACA	2302
Qy	90	GluTyrProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPro	109
Db	2303	GAGTGGCCCTACTTTGTGGTGGGAACAGTCTGTGCAATGGCAATGGAGCCCTCCAGCG	2362
Qy	110	ValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThr	129

Db	2363	GCTTCTCCATCATCTCTGCTGAGATGATAGCTATCTTTGGCCCTGGGATGACGCAGTG	2422
Qy	130	LeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPhe	149
Db	2423	AAGCAGCAAAAGTGTAACATGTTCTCCCTGCTCTTCTTGGCCCTAGGAGTCTCTCTCTC	2482
Qy	150	ValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArg	169
Db	2483	TTTACTTTCTTCTTCCAGGCTTTCACGTTTGGGAAAGCTGGAGAGATCCTCACCAAGG	2542
Qy	170	LeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGluLys	189
Db	2543	CTCCGGTCCATGGCTTTAAAGCATGCTAAGCAGACATGAGCTGGTTGTATGATCAT	2602
Qy	190	GluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGly	209
Db	2603	AAAAACAGTACTGGAGCATTCTTACAAGATCTGCCACAGATGTCGCAAGTCCAAGGA	2662
Qy	210	AlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerVal	229
Db	2663	GCCACGGGAACCAAGTTGGCTTTAATTGCAACAGCAACACGCAACCTTGGAAACCGGTATT	2722
Qy	230	IleIleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProVal	249
Db	2723	ATTATATCATTTATTTACGTTGGCACTGACACTTCTGCTGTTATCGTTGTTCCATTC	2782
Qy	250	LeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLys	269
Db	2783	ATTGCTGTAGCAGGAATGTTGAAATGTTGGTGGCAATGTCGCAAGAGAGATAAA	2842
Qy	270	GlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIle	289
Db	2843	AAGAAATGGAAGCTGCTGGAAGATGTCACACAGAGCAATAGAAAAATATTTCGAACCTGT	2902
Qy	290	ValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGln	309
Db	2903	GTATCTTGTGACCAAGAAAGAAATTTGAGTCAATGATGTTGAAATTCATGCGACCT	2962
Qy	310	HisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAla	329
Db	2963	TACAGGAATTCGGTGGGAGGACACATCTACGGCATCTTTTAGCATCTCCCAAGCA	3022
Qy	330	PheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGly	349
Db	3023	TTCATGTATTTTCTTATGCTGGCTGTTTTCGATTGCTTCTTACCTAATTTGTGAATCGA	3082
Qy	350	ArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyValAlaMetAla	369
Db	3083	CATATGGCTTCAAGATGTCATTTCTGCTCTTTTCTGCATTTGCTTGGCGCGGTGCT	3142
Qy	370	IleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHis	389
Db	3143	CTAGACACGCGCAGCTCATTTGCTCCGACATATGAAAAGCCAAAGCTGCTGTCAGCATAC	3202
Qy	390	LeuPheAlaLeuLeuGluLysLysPheAsnIleAspSerArgSerGlnGluGlyLysLys	409
Db	3203	TTGTTTCAAGCTGTTTGAAGAACCACTCTGATTGACAGCTACAGTGGAGAGGGGTGGG	3262
Qy	410	ProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArg	429
Db	3263	CCTGATAAGTTTGAAGGAAGCGTGACATTTAATGAAGTCGTGTTCACTATATCCACCCGG	3322
Qy	430	ProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAla	449
Db	3323	GCCAACTGCCAGTCTTCAGGGCTGAGCCTTGAGGTGAAGAGGGCCACAGCGTGGCC	3382
Qy	450	PheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyr	469
Db	3383	CTGTGGGCGACAGTGGCTGCGGGAAGAGCACAGTGGTCCAGCTGCTCGAGCGCTTCTAT	3442
Qy	470	AspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGln	489
Db	3443	GACCCCATGGTGGATCATGCTCTTAGATGTTCAAGAAAGCAAGAAATCAATGTCAG	3502


```
Qy 127 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 146
   :::::::::::::::::::::
Db 2365 CAGAAAGTTTTCGAGAAAGAACCTATATCTCACTGCTGTTTTAGCACCTGGGATC 2424

Qy 147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 166
   :::::::::::::::::::::
Db 2425 ATTTCCTTTTACTCTCTTGTTCAGGGTTTTCGATTTGGCAAGCTGGAGAAATCTT 2484

Qy 167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaIlePhe 186
   :::::::::::::::::::::
Db 2485 ACAATGAAGCTTCGATTCATGGCATTTAAAGCAATGCTGAGACAGGACATGGCCCTGGTTT 2544

Qy 187 AspGluLysGluAsnSerThrGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206
   :::::::::::::::::::::
Db 2545 GATGATCTAAAAATAGCACTGGAGCACTCAACAAGATGCTTAATAGTCCTCAAA 2604

Qy 207 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226
   :::::::::::::::::::::
Db 2605 GTAAAGGAGCAACTGGTGTGAGATTGGCATTTAATGTGCCAAACATAGCTAAACCTGGGA 2664

Qy 227 LeuSerValIleIleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIle 246
   :::::::::::::::::::::
Db 2665 ACTGGAATATATATCTTTAGTTTATGGCTGGCAGCTGACCCCTGTACTCTTAGCTGTT 2724

Qy 247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266
   :::::::::::::::::::::
Db 2725 GTACCAATCATTTGCAGTGGCAGGAATGATTGAATGAAGATGTTGGCTGGCGCAAAA 2784

Qy 267 LysAspLysGlnGluLeuLysHisIleGlyValIleAlaThrGluAlaLeuGluAsnIle 286
   :::::::::::::::::::::
Db 2785 AAAGATAAAATAGAGCTGGAAGCTGCAGGAAGATTTGCCACAGAAAGCCATAGAAAATATT 2844

Qy 287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306
   :::::::::::::::::::::
Db 2845 CGAACTGTGTGCTAGTTGACTCGGGAAGAAAGATTTTGACTCATGTACGGAGAACATTG 2904

Qy 307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326
   :::::::::::::::::::::
Db 2905 CTTGTACCATACAGAAATTTGTGTAAGAAGCGGCACATATTGGGTTTGTGTTTTCCTT 2964

Qy 327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyValaTyrIleu 346
   :::::::::::::::::::::
Db 2965 TCACAAGCAATGATGTTTTTACCTATCTGCTGCTGTTTCCGCTTTGGAGCCCTATCTGTG 3024

Qy 347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366
   :::::::::::::::::::::
Db 3025 GTTAATGACACATAGACTAGACAAAACCTGTGTTTTAGTGTGTTTCAGCTGTGTATTGCT 3084

Qy 367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386
   :::::::::::::::::::::
Db 3085 GCTATGGCATTAGGACAAACACAGCTCTTTTGCCACGAGACTATGCCAAAGCCAAAGATATCA 3144

Qy 387 AlaAlaHisLeuPheAlaLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 406
   :::::::::::::::::::::
Db 3145 GCAGCCCACTTGTGTTGCTGTTTAAATAGAGTACCTCCAAATAGACAGATTACCGCAGGAT 3204

Qy 407 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 426
   :::::::::::::::::::::
Db 3205 GGAGAGAACTTGAAATAATTTGGGGGAACACAAAGATCAAAAGATGTGAATAATTAACATAT 3264

Qy 427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446
   :::::::::::::::::::::
Db 3265 CCAACCCGACAGAGGTCAAAATCTCCAGGCTTGAATCTAGCAGTAGAGAAAAGAGAA 3324

Qy 447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuGln 466
   :::::::::::::::::::::
Db 3325 ACCTTGGCCCTTGTGTTGAGCAGTGGATGTGGAAGAGACACTGTTGTTTCAGCTACTTGAG 3384

Qy 467 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 486
   :::::::::::::::::::::
Db 3385 AGATTTTATGATCCCACTCAGTCGAGAAATTTGTTGTATGATGATGATGATGCAAAAGACATTA 3444

Qy 487 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506
```

```
Db 3445 AATATCCAGTGGCTGAGATCTTCACATCGGTATCTCTCACAAAGAGCCAATCTGTTCGAC 3504

Qy 507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 526
   :::::::::::::::::::::
Db 3505 TTCACCTCGCTGAACATCGCTACGGGACAAACAGTCGGGAGGTGTCCCATGAGGAA 3564

Qy 527 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGlyGlyLeuProGluLys 546
   :::::::::::::::::::::
Db 3565 ATCATTAGTCAGCGAAAGCAGCCAGTATTCTATTCCTTCATTGACTCTCTGCCAGAAAAA 3624

Qy 547 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlnLysGlnLysGlnArgLeu 566
   :::::::::::::::::::::
Db 3625 TACAATACTCGTGTAGGACAAAGGGAACACAGCTTCTTGGTGGTCAGAAACAAACGATATT 3684

Qy 567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 586
   :::::::::::::::::::::
Db 3685 GCTATTGCTGAGCTCTTATAGTAAAGCCAGATTCTGTCTACTGGATGAAGCTACATCT 3744

Qy 587 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606
   :::::::::::::::::::::
Db 3745 GCCTTGGACACAGAAAGCGAAAGATTGTCCAGGAAGCACTGGATAAAGCCAGAGAAGGT 3804

Qy 607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626
   :::::::::::::::::::::
Db 3805 CGCACCTGCATCGTAGCTCACCGCTGTGCCACCATCCAGAATGCTGCAAGAAGATTGCT 3864

Qy 627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 646
   :::::::::::::::::::::
Db 3865 GTGATCCGAAGCGAAGGTCATAGACAGGACTCACCAGAGCTCTCGCTGGTGAAGAAAA 3924

Qy 647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657
   :::::::::::::::::::::
Db 3925 GGCTTCTACTACTCGCTGGTCAATGTTCAAAAGT 3957

RESULT 15
OAU78609 OAU78609 3858 bp mRNA linear MAM 11-OCT-2001
LOCUS Ovis aries multidrug resistance protein-1 (mdr1) mRNA, complete
DEFINITION cds.
ACCESSION U78609
VERSION U78609.1 GI:2149086
KEYWORDS Ovis aries.
SOURCE Ovis aries
ORGANISM Ovis aries
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Caprinae; Ovis.
REFERENCE 1 (bases 1 to 3858)
AUTHORS Longley,M., Phua,S.H., van Stijn,T.C. and Crawford,A.M.
TITLE Isolation and mapping of the first ruminant multidrug resistance
genes
JOURNAL Anim. Genet. 30 (3), 207-210 (1999)
MEDLINE 99371931
PUBMED 10442984
REFERENCE 2 (bases 1 to 3858)
AUTHORS Longley,M. and Crawford,A.M.
TITLE Direct Submission
JOURNAL Submitted (18-NOV-1996) Biochemistry Department, University of
Otago, Dunedin, New Zealand
FEATURES
source
1. .3858
/organism="Ovis aries"
/db_xref="taxon:9940"
gene
1. .3858
/gene="mdr1"
CDS
1. .3858
/gene="mdr1"
/codon_start=1
/product="multidrug resistance protein-1"
/protein_id="AAB58489.1"
/db_xref="GI:2149087"
/translation="MDLEGRNGRAAGNFKRDKRFFSKDKKRPVSTFTWTF"
```



```
Qy 545 GluLysTyrAenThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGln 564
Db 3502 GACAAATACAAACACCAGAGTAGGGACAAAGAACTCAGCTCTCTGGTGGCCAGAAACAG 3561
Qy 565 ArgLeuAlaIleAlaArgAlaLeuGlnLysProLysIleLeuLeuAspGluAla 584
Db 3562 CGCATTTGCCATAGCTCGCGCCCTTGTAGACAGCCTCACATCTTGCTTCTGGATGAAGCT 3621
Qy 585 ThrSerAlaLeuAspAsnSerGluLysValValGlnHisAlaLeuAspLysAlaArg 604
Db 3622 ACCTCGGCTCTGGACACCGAAAGTGAAAGGTGGTCCAGAGCCCTGGACAAAGCCCGA 3681
Qy 605 ThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeu 624
Db 3682 GAGGACCGACCTGCATCGCTCACCCTGCTCCACCATCCAGAACGCAGACTTG 3741
Qy 625 IleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuArg 644
Db 3742 ATAGTGTGTTCAGAAATGGCAGAAATCAAGAGCATGGCAGACACACACAGCTGCTGGCG 3801
Qy 645 AsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 657
Db 3802 CAGAAAGGCATCTATTCCACCATGGTGTCTCAGGCT 3840
```

Search completed: March 31, 2003, 02:07:24
Job time : 8998.97 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 30, 2003, 01:20:01 ; Search time 644.062 Seconds
(without alignments)
2304.231 Million cell updates/sec

Title: US-09-873-409-1

Perfect score: 3223

Sequence: 1 MIAEKGAHAEIMAKRGDLYS.....QELLNRNDIYFKLVNAQSVQ 659

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_epool/US09873409/runat_27032003_115416_19216/app_query_fasta_1.7544
-DB=N_Geneseq_101002 -Qfmt=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09873409 @CGN 1.1 4247 @runat_27032003_115416_19216 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_WMAP -LARGQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_101002.*
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	3163	95.2	3699	24	ABK83223 Human transporter
2	1924.5	57.9	3924	21	AA294742 Human ATP binding
3	1924.5	57.9	3924	21	AA288974 Human MDR-3 DNA.
4	1924.5	57.9	3924	24	ABN95801 Gene #2299 used to
5	1882.5	56.7	3912	24	ABK63653 Rat sequence diffe
6	1882	56.6	4369	21	AA252047 Rat multidrug resi
7	1882	56.6	4425	21	AA252048 Rat multidrug resi
8	1873	56.4	4186	22	AAF86127 Cynomolgous monke
9	1873	56.4	4195	22	AAF86128 Cynomolgous monke
10	1872	56.3	2726	15	AA070907 Multidrug-resistan
11	1872	56.3	2726	15	AA070916 Multidrug-resistan
12	1872	56.3	2726	18	AA743322 Multidrug resistan
13	1872	56.3	3840	24	ABL91687 Human polynucleoti
14	1872	56.3	3860	21	AA249332 Human wild-type mu
15	1872	56.3	3860	21	AA249333 Human BCRP DNA rel
16	1872	56.3	3860	24	ABA94365 Human BCRP DNA rel
17	1872	56.3	3860	24	ABA94366 Human BCRP DNA rel
18	1872	56.3	3988	21	AA288973 Human MDR-1 DNA..
19	1872	56.3	4264	19	AAV66533 Mutated human P-gl
20	1872	56.3	4264	19	AAV66534 Mutated human P-gl
21	1872	56.3	4349	22	AAH57442 Human intestine ce
22	1872	56.3	4646	21	AA294738 Human ATP binding
23	1872	56.3	4646	24	AA238994 Human mdr1 Gene.
24	1872	56.3	4646	24	ABL68592 Kidney cancer rela
25	1872	56.3	4646	24	ABL68880 Kidney cancer rela
26	1872	56.3	4669	8	AA70752 Sequence of human
27	1872	56.3	6505	17	AA213394 Hybrid vector pSF-
28	1872	56.3	8630	21	AA224041 Retroviral M4 mdr-
29	1872	56.3	8630	21	AA224042 Retroviral vector
30	1871	56.3	4788	21	AA249335 Murine multidrug r
31	1871	56.3	4788	21	AA249336 Mouse BCRP DNA rel
32	1870	56.3	4279	22	AA03488 Dog P-glycoprotein
33	1870	56.3	4279	22	AA03488 Dog P-glycoprotein
34	1870	56.3	4279	22	AA03504 Dog P-glycoprotein
35	1870	56.3	4279	22	AA03505 Dog P-glycoprotein
36	1869	56.2	4279	22	AA03506 Dog P-glycoprotein
37	1869	56.2	4378	11	AAQ4522 Multidrug Resistan
38	1869	56.2	4669	14	AAQ52726 Sequence of human
39	1861	56.0	4646	15	AAQ72872 Human multidrug re
40	1855	55.8	4669	19	AAV32645 Human P glycoprote
41	1855	55.8	4669	24	ABK52041 cDNA encoding huma
42	1849.5	55.7	4233	21	AA290198 Rat mdr1b2 (multis
43	1849.5	55.7	4233	22	AAF27498 Rat mdr1b2 multidr
44	1844.5	55.5	4189	21	AA249334 Murine multidrug r
45	1844.5	55.5	4189	24	ABA94367 Mouse BCRP DNA rel

ALIGNMENTS

RESULT 1
ABK83223
ID ABK83223 standard; cDNA; 3699 BP.
XX
AC ABK83223;
XX
DT 27-AUG-2002 (first entry)
XX
DE Human transporter and ion channel, TRICH14, Incyte ID 7472030CBL, cDNA.
XX
KW Human; ss; gene; transporter and ion channel; TRICH; transport disorder;
KW neurological disorder; muscle disorder; immunological disorder; cancer;
KW scleroderma; systemic lupus erythematosus; allergy; leukaemia;
KW cell proliferative disorder; cervical cancer; breast cancer;
KW neurodegenerative disorder; Parkinson's disease; Alzheimer's disease;
KW myotonic dystrophy; catatonias; endocrine disorder; diabetes;
KW Grave's disease; gastrointestinal disorder; Crohn's disease;
KW renal disorder; Good pasture's syndrome; viral infection; cirrhosis;

Qy 341 PheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPhe 360
Db 2740 TTTGGAGCCTATTAAATCAAGCTGGAGCAATGACCCAGAGGGCATGTTTCATAGTTTTT 2799
Qy 361 ThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyr 380
Db 2800 ACTGCAATTCATATGGAGCTATGGCCATCGGAAACGCTCGTTTGGCTCCCTGAATAT 2859
Qy 381 SerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIle 400
Db 2860 TCCAAAGCCAAATCGGGGCTGGCGATCTGTTGGCTTTGGTGGAAAGAACCAATATA 2919
Qy 401 AspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArg 420
Db 2920 GACAGCCGACGTCAAGAAAGGCAAAAGCCAGACACATGTGAAGGCAATTTAGAGTTTCCA 2979
Qy 421 GluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeu 440
Db 2980 GAAGTCTCTTTCTTATCCATGTGCGCCAGATGTTTTCATCTCCGCTGGCTTATCCCTC 3039
Qy 441 SerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThr 460
Db 3040 AGTATTGACCGAGGAACAGTAGCATTTGTGGGAGCAGCGGCTGTGGGAAAGCACT 3099
Qy 461 SerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGly 480
Db 3100 TCTGTCAACTTCTGCAGAGACTTTATGACCCCGTGCAAGGACAGTCTGTTGATGGT 3159
Qy 481 ValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGln 500
Db 3160 GTGGATGCAAAAGAAATTGAATGTACAGTGGCTCCGTTCCCAATAGCAATCGTTCCCTCAA 3219
Qy 501 GluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArg 520
Db 3220 GAGCCTGCTCTTCACTGCAGCATTTCTGAGAACATCGCTTGGTGGCAACAGCCGT 3279
Qy 521 ValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIle 540
Db 3280 GTGGTGCCATTAGATGAGATCAAGAGCCGCAATGACGCAAAATATCCATTTCTTTAT 3339
Qy 541 GluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGly 560
Db 3340 GAAAGTCTCCCTGAGAAATACACACACAAGTTGGACTGAAAGGAGCACAGCTTTCTGGC 3399
Qy 561 GlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuGlnLysProLysIleLeuLeu 580
Db 3400 GGCCAGAAACAAAGACTAGCTATTGCAAGGGCTCTTCTCCAAAAACCCAAATTTTATTTG 3459
Qy 581 LeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeu 600
Db 3460 TTGGATGAGGCCACTTCAGCCCTCGATTAATGACATGAGAGGTGTTTCAGCATGCCCTT 3519
Qy 601 AspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGln 620
Db 3520 GATAAGCCAGGACGAGGAGGACATGCTAGTGGTCACTCACAGGCTCTCTGCAATTCAG 3579
Qy 621 AsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGln 640
Db 3580 AACGCGAGATTTGATAGTGGTCTGACCAATGGAAGATAAAGGAACAAGGAACTCATCAA 3639
Qy 641 GluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 659
Db 3640 GAGCTCTCTGAGAAATCGACACATATATTTTAAGTTAGTGAATGCACAGTCAGTCGAG 3696

RESULT 2
AAZ94742
ID AAZ94742 standard; cDNA; 3924 BP.
XX
AC AAZ94742;
XX
DT 01-AUG-2000 (first entry)
XX

DE Human ATP binding cassette ABCB4 (MDR3) cDNA.
XX ABCB4; ATP binding cassette; human; cholesterol; lipid disorder;
KW atherosclerosis; lipid disorder; dyslipidemia; psoriasis;
KW lupus erythematosus; diagnosis; gene therapy; MDR3;
KW multidrug resistance; chromosome 7q21; ss.
XX
OS Homo sapiens.
XX
XX WO200018912-A2.
XX
XX 06-APR-2000.
XX
XX 21-SEP-1999; 99WO-EP06991.
XX
XX 25-SEP-1998; 98US-0101706.
XX
XX (FARB) BAYER AG.
XX
XX Schmitz G, Klucken J;
XX
XX WPI; 2000-293151/25.
XX
XX Adenosine triphosphate binding proteins useful for identifying agents
PT for treating atherosclerosis and other inflammatory disorders -
PT
XX
XX Claim 9; Page 117-119; 154pp; English.
XX
XX The present sequence is that of human ATP binding cassette
CC subfamily B protein ABCB4 cDNA. The cDNA was identified using a
CC differential display method in which monocytes from peripheral
CC blood were subjected to macrophage differentiation and cholesterol
CC loading with acetylated low density lipoproteins and subsequent
CC deloading with high density lipoprotein (HDL3) to identify
CC cholesterol sensitive genes. The gene maps to chromosome 7q21
CC and is also termed MDR3 (multidrug resistance). The invention
CC provides cholesterol-sensitive ABC genes (see AA294734-63). These
CC genes, and polypeptides encoded by them, can be used for diagnostic
CC and therapeutic applications, and for biochemical or cell-based
CC assays to screen for pharmacologically active modulator compounds
CC useful for the treatment of lipid disorders, atherosclerosis or
CC other inflammatory diseases such as psoriasis and lupus
CC erythematosus.
XX
XX Sequence 3924 BP; 1145 A; 790 C; 977 G; 1012 T; 0 other;
SQ

Alignment Scores:
Pred. No.: 5,43e-187 Length: 3924
Score: 1924.50 Matches: 372
Percent Similarity: 74.10% Conservative: 123
Best Local Similarity: 55.69% Mismatches: 162
Query Match: 57.91% Indels: 11
DB: 21 Gaps: 3

US-09-873-409-1 (1-659) x AAZ94742 (1-3924)
Qy 1 MetLeuAlaGluLysGlyValaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
Db 1851 GTAATTGGAGCAGGAGGACGACAGCACTGATGAGAGGAGGAGGGGTGTACTTCAA 1910
Qy 21 LeuVal-----MetSerGlnAspIleLysLysAlaAsp 31
Db 1911 CTTGTCAACATGCAGACATCAGGAAGCCAGATCCAGTCAGAAGAATTTGAACATAATGAT 1970
Qy 32 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis 51
Db 1971 GAAAGGCTGCCACTAGAAATCGCCCAATCGCTGGAATCTCGCTATTATAGGCAATCT 2030
Qy 52 SerValLysSerIleLys---SerAspPheIleAspLysAla---GluGluSerThrGln 69
Db 2031 ACTCAGAAAACCTTAAATAATTCACAAATGTGTGAGAGAGCCTGTGTAACCCGAT 2090
Qy 70 SerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysPro 89


```
Db 2091 GCATTGAAGCAAAATGTCACCAGTGTCTTTCTGAAGGTCTGAAACTGAATAAACA 2150
Qy 90 GluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAenGlyThrValHisPro 109
Db 2151 GAATGGCCCTACTTGTGGGGAAAGAGTATGCGCAATGCGCAATGGGGGCTTCAGCCG 2210
Qy 110 ValPheSerIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThr 129
Db 2211 GCATTTTCAGTCATATCTCAGAGATCATAGCGATTTTGGACAGCGCATGTCAGTG 2270
Qy 130 LeuLysHisAspAlaGluLeuIleThrSerMetIlePheValIleLeuGlyValIleCysPhe 149
Db 2271 AAGCAGCAGAAAGTGCACATATTTCTTTGATTTCTTCTTCTTCTTCTTCTTCTTCTT 2330
Qy 150 ValSerThrPheMetGlnGlyLeuPheThrGlyArgAlaGlyGluIleLeuThrMetArg 169
Db 2331 TTTACTTTCTCTCTCAGGGTTTCAGTGTGGGAAGCTGGGAGATCTCCACGAGA 2390
Qy 170 LeuArgHisLeuAlaPheLysAlaMetLeuThrGlnAspIleAlaTrpPheAspGluLys 189
Db 2391 CTGCGTCAATGCTTTAAAGCAATGCTAAGACAGCATGAGCTGGTTGATGACCAT 2450
Qy 190 GluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGly 209
Db 2451 AAAACAGTACTGGTGCACTTTTACAAGACTTGCACAGATGCTGCCCAAGTCCAAGGA 2510
Qy 210 AlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerVal 229
Db 2511 GCCACAGGAACCAAGTGGCTTAAATTTGACACAGATATAGCTTAACCTTGGAAT 2570
Qy 230 IleIleSerPheIleThrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProVal 249
Db 2571 ATCATATCATTTATCTACGGTTGGCAGTTAAACCTTATGCTATTAGCAGTTGTTCCAAT 2630
Qy 250 LeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLys 269
Db 2631 ATTGCTGTGTGAGGAATTTGTGAATTTGAAATTTGTTGGCTGGAATGCCAAAAGAGATAA 2690
Qy 270 GlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIle 289
Db 2691 AAGAACTGGAAGCTGCTGGAAGAGTTGAACAGAGGCAATAGAAAATATTAGGACAGTT 2750
Qy 290 ValSerLeuThrArgGluLysAlaPheGluGlnMetThrGluGluMetLeuGlnThrGln 309
Db 2751 GTGCTTTGACCCAGGAAGAAAATTTGAATCAATGATGTTGAAAATTTGATGACCT 2810
Qy 310 HisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAla 329
Db 2811 TACAGGAATTTCTGCAGAAAGGCACACATCTATGGAATTTACTTTTAGTATCTCACAGCA 2870
Qy 330 PheIleThrPheAlaThrAlaAlaGlyPheArgPheGlyAlaThrLeuIleGlnAlaGly 349
Db 2871 TTTATGATATTTTCTCTATCCCGTGTGTTTTCGATTTGGTGCAATCTCATTTGGAATGGA 2930
Qy 350 ArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaIleAlaMetAla 369
Db 2931 CATATCGGCTTCAGATGTTATTCGGTGTTTCTGCAATTTGATTTGGTCAGTGGCT 2990
Qy 370 IleGlyLysThrLeuValLeuAlaProGluThrSerLysAlaLysSerGlyAlaAlaHis 389
Db 2991 CTAGGACATGCCAGTTTCATTTGCTCCAGACTATGCTAAAGCTGTCGAGGCCAC 3050
Qy 390 LeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluLysLys 409
Db 3051 TTATTCATCTGTTTGAAGACAAACCTCTGATTCAGAGCTACAGCTAAGAGGGGCTGAAG 3110
Qy 410 ProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheThrProCysArg 429
Db 3111 CTGATAAATTTGAAGAAATATAACATTTAATGAAGTCTGTTCAACTATCCACCCGA 3170
Qy 430 ProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAla 449
```

```
Db 3171 GCAAAACGTGCCAGTCTTTCAGGGGCTGAGCCTGGAGGTGAAGAAAGCCAGACACTAGCC 3230
Qy 450 PheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyr 469
Db 3231 CTGGTGGGACAGAGTGGCTGTGGGAAGACAGCGTGGTCCAGCTCTTGAGCGGTCTTAC 3290
Qy 470 AspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGln 489
Db 3291 GACCCTTTGGGGGGACAGTGTCTTCGATGTGTCAAGAAGCAAGAAATCAATGTCACG 3350
Qy 490 TrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIle 509
Db 3351 TGGCTCAGAGCTCAATCGGAATCGTCTCAGGAGCCTATCTCTATTTGACTGACGACT 3410
Qy 510 AlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGlu 529
Db 3411 GCGGAGAATATTGCTTATGGAGACCAACGCGGTGTATCACAGATGAATTTGTGAGT 3470
Qy 530 AlaAlaAsnAlaAlaAsnIleHisSerPheIleGlyLeuProGluLysTyrAsnThr 549
Db 3471 GCAGCCAAAGCTGCCCAACATACATCTTTTCATCGAGACGTTTACCCACAAATATGAACA 3530
Qy 550 GlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAla 569
Db 3531 AGAGTGGGAGATGAAGGGACTCAGCTCTCAGGAGGTCAAAAACAGAGGATTGCTATTGCC 3590
Qy 570 ArgAlaLeuLeuGlnLysProLysIleLeuLeuAspGluAlaThrSerAlaLeuAsp 589
Db 3591 CGAGCCCTCATCAGCAACCTCAAAATCTCTCTTGGATGAAGCTACATCAGCTCTGGAT 3650
Qy 590 AsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCys 609
Db 3651 ACTGAAAGTGAAGAGTTGTCCAGAGCCCTGGCAAAAGCCAGAGAGGCCGACCTGC 3710
Qy 610 LeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHis 629
Db 3711 ATTTGATTTGCTCACCGCTGTCCACCATCTCAGAAATGCAGACTTAATAGTGGTGTTCAG 3770
Qy 630 AsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyr 649
Db 3771 AATGGAGAGATCAGAGCATGGCAGCATGACGAGCTGCTGGGCACAGAAAGGCATCTAT 3830
Qy 650 PheLysLeuValAsnAlaGlnSer 657
Db 3831 TTTTCAATGTCAGTGTCCAGCT 3854
RESULT 3
AAZ88974
ID AAZ88974 standard; DNA; 3924 BP.
XX AC AAZ88974;
XX DT 30-MAY-2000 (first entry)
XX DE Human MDR-3 DNA.
XX KW Phosphatidylcholine; treatment; disease; multidrug resistance; MDR-3;
XX KW stomach; colon; ulcerative colitis; pouchitis; large intestine; human;
XX KW inflammation; Crohn's disease; diverse colitis; carcinoma; gene therapy;
XX KW ileum; anti-ulcer; anti-inflammatory; cytostatic; ss.
XX OS Homo sapiens.
XX PN WO200007577-A2.
XX PD 17-FEB-2000.
XX PF 06-AUG-1999; 99WO-DE02426.
XX PR 06-AUG-1998; 98DE-1035526.
XX PR 15-DEC-1998; 98DE-1057750.
XX PA (STRE/) STREMMEL W.
```


XX Stremmel W;
XX WPI; 2000-195439/17.
XX
XX Using phosphatidylcholine having mucous membrane protective activity to
XX prevent or treat large intestinal diseases, such as ulcerative colitis
XX or pouchitis -
XX
XX Disclosure; Page 17-22; 22pp; German.
XX
XX This invention describes a novel medicament containing a therapeutically
XX effective amount of phosphatidylcholine to treat diseases, where
XX phosphatidylcholine has an advantageous mucous membrane protective
XX activity in the large intestine. The invention describes (1) a method
XX to regulate presence or absence of MDR(multidrug resistance)-3
XX transcription in a sample from patients with suspicion of stomach and
XX colon diseases, where treatment with phosphatidylcholine is indicated;
XX (2) a diagnostic test, to detect ulcerative colitis, pouchitis, large
XX intestinal inflammation, Crohn's disease, diverse colitis, infectious
XX enteritis/colitis, inflammation through X-ray treatment, antibiotics,
XX chemotherapeutics, drugs or chemicals or large intestinal carcinoma,
XX contains MDR3-specific primers to detect mutations in the MDR3 gene
XX family and their interactions partners (in particular transcription
XX factors); and (3) a method to manufacture a medicament for local gene
XX therapy in the ileum with the ideal MDR3-analogous genes for patients
XX with missing or decreased expression of MDR3-analogous proteins. The
XX products of the invention have anti-ulcer, anti-inflammatory and
XX cytosstatic activity. The orally administered delayed-release form of
XX phosphatidylcholine prevents premature resorption and provides for
XX targeted release in the lower section of the small or large intestine.
XX This sequence encodes the human MDR-3 protein described in the method of
XX the invention.
XX
SQ Sequence 3924 BP; 1145 A; 790 C; 977 G; 1012 T; 0 other;

Alignment Scores:
Pred. No.: 5,43e-187 Length: 3924
Score: 1924.50 Matches: 372
Percent Similarity: 74.10% Conservative: 123
Best Local Similarity: 55.69% Mismatches: 162
Query Match: 57.91% Indels: 11
DB: 21 Gaps: 3

US-09-873-409-1 (1-659) x AAZ88974 (1-3924)

Qy 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrSer 20
Db 1851 GTAATTGTGGAGCAAGGAAGCCACAGCGAACTGTGAAGAGGAAGGGGTGACTTCAA 1910
Qy 21 LeuVal-----MetSerGlnAspIleLysLysAlaAsp 31
Db 1911 CTGTGCAACATCGACATCAGGAGCCAGATCCAGTCAGAGAAATTTGAATCAATGAT 1970
Qy 32 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis 51
Db 1971 GAAAGGCGTCCACTAGATGCCCCCAATGGCTGGAATCTCGCTATTAGGCATTCT 2030
Qy 52 SerValLysSerIleLys---SerAspPheIleAspLysAla---GluGluSerThrGln 69
Db 2031 ACTCAGAAAAACCTTAAAAAATTCACAAATGTGTCTCAGAGAGCGCTTGTATGGAAAACCGAT 2090
Qy 70 SerLysGluLeuSerLeuProGluValSerLeuLysIleLeuLysLeuAsnLysPro 89
Db 2091 GGACTTGAAGCAAAATGTGCCACAGTGTCTTTCTGAAGGTCTGAAACTGAATAAACA 2150
Qy 90 GluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPro 109
Db 2151 GAATGGCCCTACTTTGTCTGGGACAGATATGTGCCAATGGGGGGCTTCAGCG 2210
Qy 110 ValPheSerIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThr 129
Db 2211 GCATTTTTCAGTCATATCTCAGAGATCATAGCGATTTTGGACCGAGCGATGATCGATG 2270

Qy 130 LeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPhe 149
Db 2271 AAGCAGCAGAAAGTGCACAATATCTCTTTGATTTCTTATTCTCGGAATATTCTTTT 2330
Qy 150 ValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArg 169
Db 2331 TTTACTTTCTTCTTCAGGGTTTCAAGTTTGGGAAAGCTGCGGAGATCTCTCACCAGAAGA 2390
Qy 170 LeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLys 189
Db 2391 CTGGGCTCAATGGCTTTTAAAGCAATGCTAAGACAGCAGCATGAGCTGGTTGATGACCAT 2450
Qy 190 GluAsnSerThrGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGly 209
Db 2451 AAAAACAGTACTGCTGCACTTTCTACAAGACTTGCACAGATGCTGCCCAAGTCCAAGGA 2510
Qy 210 AlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerVal 229
Db 2511 GCCCAGGAACCGGTTGGCTTTAATTGCAACAGAAATATAGCTAACCTTGGAACTGGTATT 2570
Qy 230 IleIleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProVal 249
Db 2571 ATCATATCATTTATCTAGCTTGGCAGTTAACCTTATGCTATTAGCAGTTGTTCCAATT 2630
Qy 250 LeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLys 269
Db 2631 ATTGCTGTGTGAGAAATTTGTAATGAATGTTGTGGCTGGGAAATGCCAAAGAGATAAA 2690
Qy 270 GlnGluLeuLysHisAlaGlyIleAlaThrGluAlaLeuGluAsnIleArgThrIle 289
Db 2691 AAAGAACTGGAGCTGCTGGAAAGATTGCCACAGAGCAATAGAAAAATATTAGGACAGTT 2750
Qy 290 ValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGln 309
Db 2751 GTGCTTTGACCCAGGAAGAAATTTGAATCAATGATGTTGAAAAATTTGTATGGACCT 2810
Qy 310 HisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAla 329
Db 2811 TACAGGAATTTCTGTGCAGAGGCACACATCTATGGAATTTACTTTTAGTATCTCACAAGCA 2870
Qy 330 PheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGly 349
Db 2871 TTTATGATATTTTCTATGCGGGTTGTTTTCGATTTGTGTCATATCTCATTTGTGAATGGA 2930
Qy 350 ArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAla 369
Db 2931 CATATGCCCTTCAGAGATGTTATCTGGTGTTTCTGCAATTTGTTTGGTGCAGTGCCT 2990
Qy 370 IleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHis 389
Db 2991 CTAGACATGCCAGTTTCATTTGCTCCAGCTATGCTAAAGCTAAGCTGTCTCAGCCCCAC 3050
Qy 390 LeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLys 409
Db 3051 TTTATTCATGCTGTTTGAAGAGACACCTCTGATTTGACACTACAGTGAAGAGGGGTGAAG 3110
Qy 410 ProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArg 429
Db 3111 CTTGATAAATTTGAAGGAATATAACATTTAATGAAGTGTGTTCACTATATCCACCCGA 3170
Qy 430 ProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAla 449
Db 3171 GCAAACGTCAGCTCTTCAGGGCTGAGCTGGAGGTGAAGAAAGGCCAGACACTAGCC 3230
Qy 450 PheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuGlnArgLeuTyr 469
Db 3231 CTGTGGGCGCAGCTGTGGGAGAGCAGCGTGGTCCAGCTCTCTGAGCGGTTCTAC 3290
Qy 470 AspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGln 489
Db 3291 GACCCCTTGGCGGCGACAGTCTCTCGATGTTCAAGAAAGCAAGAAATCAATGTTCCAG 3350

Db 2511 GCCCAGGACACGAGTGGCTTTAATGTCACAGAAATATAGCTTAACCTTGGAACTGGTATT 2570
Qy 230 IleIleSerPheIleTyRgLyTrpGluMetThrPheLeuIleLeuSerIleAlaProVal 249
Db 2571 ATCATATCATTTATCTACGGTTGGCAGTTAACCTTATTGCTATTAGCAGTGTGTTCCAATT 2630
Qy 250 LeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLys 269
Db 2631 ATTGCTGTCTCAGGAATTTGTTGAAATGAATTTGGTGGGAAATGCCCAAGAGATATA 2690
Qy 270 GlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIle 289
Db 2691 AAGAAGCTGGAAAGCTGCTGGAAGGATTGCAAGAGGCAATAGAAAATATTAGGACAGTT 2750
Qy 290 ValSerLeuThrArgGluLysAlaPheGluGlnMetTyRgGluMetLeuGlnThrGln 309
Db 2751 GTGTCTTTGACCAGCAAGAAAATTTGAATCAATGATGTTGAAAATTTGATGGACCT 2810
Qy 310 HisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyRAlaPheSerHisAla 329
Db 2811 TACAGGAATTTCTGTCAGAGGCACACATCTATGGAATTTACTTTTAGTATCTCACAAGCA 2870
Qy 330 PheIleTyRAlaAlaGlyPheArgPheGlyAlaTyRLeuIleGlnAlaGly 349
Db 2871 TTTATGATTTTCTATGCTGCGGTTGTTTCGATTTGGTGCATATCTCATTTGTAATGGA 2930
Qy 350 ArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyRAlaMetAla 369
Db 2931 CATATGCGCTTCAGAGATGTTATCTGGTGTGTTTCTGCNAATGTTATGGTGCAGTGCT 2990
Qy 370 IleGlyLysThrLeuValLeuAlaProGluTyRSerLysAlaLysSerGlyAlaAlaHis 389
Db 2991 CTAGGACATGCCAGTTTCATTTGCTCCAGACTATGCTAAAGCTAAGCTGCTCAGCCCCAC 3050
Qy 390 LeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLys 409
Db 3051 TTTATTCATGCTCTTTTGAAGACAACCTCTGATTTGACAGCTACAGTGAAGAGGGGCTGAAG 3110
Qy 410 ProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyRProCysArg 429
Db 3111 CTGTATAAATTTGAAGAAATATACATTTAATGAAGTCTGTTCAACTATCCACCCGA 3170
Qy 430 ProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAla 449
Db 3171 GCAAAAGCTGCCAGTGCTTCAGGGGCTGAGCCTGGAGGTGAAGAAAGGCCAGACATAGCC 3230
Qy 450 PheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyR 469
Db 3231 CTGTGGGCGACAGTGGCTGTGGGAAGACACGGTGGTCCAGCTCTCTGGAGCGGTTCTAC 3290
Qy 470 AspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGln 489
Db 3291 GACCCCTTGGCGGGACAGTGTCTCGATGTCAAGAGCAAGAACTCAATGTCCAG 3350
Qy 490 TrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIle 509
Db 3351 TGGCTCAGAGCTCAACTCGAATCGTGTCTCAGAGGCTTATCTTATTGACGTGCAGCAAT 3410
Qy 510 AlaGluAsnIleAlaTyRgLyAspAsnSerArgValValProLeuAspGluIleLysGlu 529
Db 3411 GCCGAGAATATTGCGCTATGGAGACAACACCGCGGTTGTATCACAGGATGAAATTTGTGAGT 3470
Qy 530 AlaAlaAsnAlaAlaAsnIleHisserPheIleGluGlyLeuProGluLysTyRAsnThr 549
Db 3471 GCAGCCAAAGCTTGCCAAACATACATCTTTTCATCGAGAGCTTACCCACAAAATATGAACA 3530
Qy 550 GlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAla 569
Db 3531 AGAGTGGGAGATAAGGGGACTCAGCTCTCAGAGGTCAAAACAGAGGATTCCTATTGCC 3590
Qy 570 ArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAsp 589

Db 3591 CGAGCCCTCATCAGACAACCTCAAAATCTCTCTGTTGATGAAGCTACATCAGCTCTGGAT 3650
Qy 590 AsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCys 609
Db 3651 ACTGAAGTGAAGAGTTGTCCAAGAGCCCTGGACAAGCCAGAGAAGCGCGCACCTGC 3710
Qy 610 LeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHis 629
Db 3711 ATTGTGATGTCTACCGCCTGTCCACCATCCAGAAATGCAGACTTAATAGTGGTGTTCAG 3770
Qy 630 AsnGlyLysIleLysGluGlnGlyThrHisGlnLeuLeuLeuArgAsnArgAspIleTyR 649
Db 3771 AATGGAGAGTCAGGAGCATGGCAGCATCAGCAGCTGCTGGCACAAAGGCATCTAT 3830
Qy 650 PheLysLeuValAsnAlaGlnSer 657
Db 3831 TTTTCAATGGTCAGTGTCCAGGCT 3854
RESULT 5
ABK63653
ID ABK63653 standard; cDNA; 3912 BP.
XX
AC ABK63653;
XX
DT 18-JUN-2002 (first entry)
XX
DE Rat sequence differentially expressed in response to a hepatotoxin #1560.
XX
KW Rat; ss; hepatotoxin; expressed sequence tag; EST; drug screening;
KW differential expression; centrilobular necrosis; steatosis.
XX
OS Rattus norvegicus.
XX
PN WO200210453-A2.
XX
PD 07-FEB-2002.
XX
PF 30-JUL-2001; 2001WO-US23872.
XX
PR 31-JUL-2000; 2000US-222040P.
PR 02-NOV-2000; 2000US-244880P.
PR 11-MAY-2001; 2001US-290029P.
PR 15-MAY-2001; 2001US-290645P.
PR 22-MAY-2001; 2001US-292336P.
PR 06-JUN-2001; 2001US-295798P.
PR 13-JUN-2001; 2001US-297457P.
PR 19-JUN-2001; 2001US-298884P.
PR 09-JUL-2001; 2001US-303459P.
XX
FA (GENE-) GENE LOGIC INC.
XX
PI Mendrick D, Porter MW, Johnson KR, Castile AL, Elashoff MR;
XX
XX WPI; 2002-241625/29.
XX
PT Predicting toxic effects of compounds or the progression of these toxic
PT effects by determining the changes in gene expression in tissues or
PT cells exposed to the toxin and comparing these to gene expression in
PT unexposed tissues or cells -
XX
PS Claim 1; Seq ID No 1560; 239pp; English.
XX
CC The invention relates to methods for predicting toxic effects of
CC compounds or the progression of these toxic effects by determining the
CC global changes in gene expression in tissues or cells exposed to the
CC toxin and comparing these to gene expression in unexposed tissues or
CC cells. Also included are methods of predicting at least one toxic
CC effect of a compound or progression of a toxic effect, preferably the
CC hepatotoxicity of a compound, comprising detecting the level of
CC expression in a tissue or cell sample exposed to the compound of two or
CC more genes listed in the specification, where differential expression of
CC the genes is indicative of at least one toxic effect or progression.
CC The method can also be used to identify an agent which modulates the

CC toxic response and predict cellular pathways that a compound modulates
 CC in a cell. The methods utilise a set of at least two probes (on a solid
 CC support in kit form), where each of the probes comprises a sequence that
 CC specifically hybridises to a gene listed in the specification, a computer
 CC system comprising a database containing information identifying a gene
 CC expression level in a tissue or cell sample exposed to a hepatotoxin of a
 CC set of genes comprising at least two genes listed in the specification,
 CC and a user interface to view the information used to present information
 CC identifying the expression level in a tissue or cell of at least one gene
 CC listed in the specification. The method is useful for elucidating global
 CC changes in gene expression and for identifying toxicity markers in
 CC tissues or cell exposed to a known toxin. The genes may be used as
 CC toxicity markers in drug screening and toxicity assays. The genes and
 CC gene expression information may be used as diagnostic markers for the
 CC prediction or identification of the physiological state of tissue or cell
 CC sample that has been exposed to a compound or agent. Hepatotoxicity
 CC is characterised by centrilobular necrosis and steatosis. The present
 CC sequence is an expressed sequence tag (EST) or cDNA derived from a gene
 CC which is differentially expressed in response to a hepatotoxic agent.

XX
 SQ Sequence 3912 BP; 1077 A; 904 C; 1017 G; 914 T; 0 other;

Alignment Scores:

Pred. No.: 1 12e-182 Length: 3912
 Score: 1882.50 Matches: 383
 Percent Similarity: 73.43% Conservative: 129
 Best Local Similarity: 54.18% Mismatches: 165
 Query Match: 56.65% Indels: 13
 DB: 24 Gaps: 2

US-09-873-409-1 (1-659) x ABK63653 (1-3912)

QY 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrSer 20
 DB 1842 GTCATCGTGGAGCAAGAGCCACAGTACGTAGTATAAGAGGAGGATCTACTTCAGA 1901
 QY 21 LeuValMetSerGln-----AspLeuLysLys 29
 DB 1902 CTTGTTAAATGATGCACATCAGAGCCAGATCCTGTCAGAGAAATTTGAAGTTGAGCTA 1961
 QY 30 AlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuPro 49
 DB 1962 AGTGATGAAGAGCTCTGGAGGTGGGCCCAATAGGCTGGTGAAGACGATATTAGG 2021
 QY 50 LeuHisSerValLysSerIleLysSerAsp-----PheIleAspLysAlaGluGluSer 67
 DB 2022 AATTCTACGAAGAAAGTCTGAAAGTTCACGGCGCATCAAAATAGGCTGGATGGAA 2081
 QY 68 ThrGlnSerLysGluLysSerLeuProGluValSerLeuLysLysIleLysLeuAsn 87
 DB 2082 ACCAATGAATTCATGCAACACGTCGCCACCATGCTCTTTTGAAGGCTTAAGACTGAAT 2141
 QY 88 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107
 DB 2142 AAACAGAGTGGCCCTACTTTGGTGGGACACTCTGTGCCATTCACACGGGGCCCTC 2201
 QY 108 HisProValPheSerIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLys 127
 DB 2202 CAGCGGCATCTCCATCATCTCTGACAGATGATAGTATCTTTGGCCCTGGGATGAC 2261
 QY 128 ThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIle 147
 DB 2262 ACAGTAAAGCAACAGAGTGAATGTTCTCGCTGGTCTTCTTGGGCTAGGAGTCCAC 2321
 QY 148 GysPheValSerThrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThr 167
 DB 2322 TCCTTTTACTTCTCTCTCGAGTTTCACATTCGGGAAAGCTGGCAGATCTCCACC 2381
 QY 168 MetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAsp 187
 DB 2382 ACNAGGCTCGGTCCATGGCTTCAAGCNAATGCTTAAGACAGGACATGAGTGGTTGAC 2441
 QY 188 GluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIle 207

DB 2442 GATCATAAAAACAGTACTGGTCCCTCTCTCAAGACTCGCCACAGACGCTCGGAGGTC 2501
 QY 208 GlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeu 227
 DB 2502 CAAGAGCCACAGAACACGAGTGGCTTTAATGTCACAGAACACAGCCACCTTGGACG 2561
 QY 228 SerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAla 247
 DB 2562 GGTATTATTATCATTTATTACGGTTGGCAACTGACACTTCTGCTTATCAGTTGT 2621
 QY 248 ProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLys 267
 DB 2622 CATTTCATTGCTGTAGCGGAATTTGAAATGAAATGTTGGTGGCAACCCCAAGAGA 2681
 QY 268 AspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArg 287
 DB 2682 GATAAAAGGAGATGGAAGCTCTGGAAGATTGCAACAGAGGCAATAGAAAATATTCG 2741
 QY 288 ThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGln 307
 DB 2742 ACTGTTGTATCTCTGACCCAGAGAGAAAATTTGAGTCAATGTATGTTGAAAAATTACAC 2801
 QY 308 ThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSer 327
 DB 2802 GGACCTTACAGGAATTCAGTCGGAGGCTCACATCTACGGCATCATTCTTACCATCTCA 2861
 QY 328 HisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGln 347
 DB 2862 CAAGCATTCATGATCTTTCTATGCTGGCTGCTTCGATTGGTCTTACCTCATTTGTG 2921
 QY 348 AlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAla 367
 DB 2922 AATGACACATCGCTTCAAGGATGCTCATCTCTGGTGTCTCAGCAATCGTGTGTTGCA 2981
 QY 368 MetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAla 387
 DB 2982 GTGGCTTAGGACATGCCAGCTCATTTGCTCCAGACTATGCAAAAGCCAAAGCTGTGCA 3041
 QY 388 AlaHisLeuPheAlaLeuLysLysProAsnIleAspSerArgSerGlnGluGly 407
 DB 3042 GCATACTTATTCAGTCTGTTGAAAGACAACTCTGATTGACAGTACAGACAGAGGA 3101
 QY 408 LysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrPro 427
 DB 3102 ATGTGCGCGGATAAGTTTGAAGAGCGTGACATTCATCAATGAAGTTGTGTTCAATTATCC 3161
 QY 428 CysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThr 447
 DB 3162 ACCCGGGCCCATGTGCCAGTGTCTCAGGGCTGAGCCTCGAGGTGAAGAGGGGAGACC 3221
 QY 448 ValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuArg 467
 DB 3222 CTGGCCCTGGTGGGACGATAGTGGCTCGCGGAAGACACCGTGTCCAGTGTCTCAGGCGC 3281
 QY 468 LeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsn 487
 DB 3282 TTCTACGACCCCATGGCCGGAACAGTGTCTCTCGATGTGTCAGGAAGCAAGAACTCAAT 3341
 QY 488 ValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCys 507
 DB 3342 GTCCAGTGGCTCCGAGCTCAACTTGGCATTTGTGCCAGGACCCATCTGTTTACTGTC 3401
 QY 508 SerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIle 527
 DB 3402 AGCATCGCCAGAACATCGCTTACGAGACAAACACCGTGTCTGTTCTCAGGATGAGATT 3461
 QY 528 LysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGlyGlyLeuProGluLysTyr 547
 DB 3462 GTGAGGGCGGCCAAGAGGCCCAACATCCACCCCTTCATTGAGACACTGCCCCCAAGTAT 3521
 QY 548 AsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAla 567

Db 2920 GTTCCCATCTTCTATAGCAGGAGTGGTTGAAATGTAATGTTGTCTGCACAAAGCGCTG 2979
Qy LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAenIle 286
Db 2980 AAGATAAGAAGAACTAGAGGTTCTGGAGAGTCTGACAGCAATGGAACACTT 3039
Qy ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306
Db 3040 CGCACTGTGCTCTTGTGACTCGGAGCAGAGTTTGAATATGATGATGCCAGACTTG 3099
Qy 307 GlnThrGlnHisArgSerThrSerLysLysAlaGlnIleLeuGlySerCysTyrAlaPhe 326
Db 3100 CAGATCCATACAGAATGCTTTGAGAAAGCGCAGCTCTTTGGGATCACTTTCTCTCTC 3159
Qy 327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 346
Db 3160 ACCAGGCCATGATGATTTCTCTATGCTGCTGTTTCGGTTGATGCTACTTTGGT 3219
Qy 347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366
Db 3220 GCACGAGAACTCATGACATTTGAAATGTTCTGTAGTATTTCTCAGCTATTGCTTGGT 3279
Qy 367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386
Db 3280 GCATCGCAGTGGGCGAGTCACTTCATTCGCTCTGACTACCGCAAGCCAAAGTCTCG 3339
Qy 387 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 406
Db 3340 GCATCCCATCATCATCAGGATCATTTGAGAAATCCCTGAGATTGACAGCTACACGCGAG 3399
Qy 407 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 426
Db 3400 GGCTTCAAGCCTAATATGTTGGAAGCAATGTGAAATTTAATGGAGTCATGTTCAACTAT 3459
Qy 427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446
Db 3460 CCCACCAGCCCAACATCCAGTGTCTCAGGGCTCAGGCTAGAGTGAAGAAAGGCGAG 3519
Qy 447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuGln 466
Db 3520 ACGTGGCTCTGTGGCAGCAGTGGCTGGGGAAGATACAGTGTCTCAGCTGCTTGAG 3579
Qy 467 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 486
Db 3580 CGCTTCTATGACCCCATCGCCGACAGTGTCTTAGATGGCAAGAAATAAAGCAACTC 3639
Qy 487 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnLeuProValLeuPheAsn 506
Db 3640 AATGTCCAGTGGCTCCGCGCCACCTGGGCATTTGTCTCCAGGAGCCCATCTCTTTGAC 3699
Qy 507 CysSerIleAlaGluAenIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 526
Db 3700 TGCAAGATCCCGAGAACATGCTTAGGAGACACAGCGGTCTGTCTCATAGGAG 3759
Qy 527 IleLysGluAlaAlaAsnAlaAlaAenIleHisSerPheIleGluGlyLeuProGluLys 546
Db 3760 ATCGTGAAGCGCCCAAGAGGCGCAACATCCACAGTTTCATCGACTCACTGCTGAGAAA 3819
Qy 547 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 566
Db 3820 TACAACACAGAGTGGGAGACAAAGGACTCAGCTCTCGGGCGGAGAGCAGCGCATC 3879
Qy 567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 586
Db 3880 GCATCGCGCGCCCTGTCAGACAGCTCATCTTACTTCTGATGAGGAGCATCA 3939
Qy 587 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606
Db 3940 GCTCTGATACGGAGAGTGAAGAGTCTGTCAGAGCGCTGGACAAAGCCAGGAGGC 3999
Qy 607 ArgThrCysLeuValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626

Db 4000 CGCACTGCATTTGTGATCGCGCACCGCTGTCCACATCCAGACGACACTTGATCGT 4059
Qy 627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnLeuLeuArgSerArg 646
Db 4060 GTGATTTCAGAACGGCAGGTCAAGGACGACGCGCACCCACGACGCTGTGCGCCAGAAA 4119
Qy 647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657
Db 4120 GGCACTATTTCGATGCTGATGCTGAGGCT 4152
RESULT 7
AAZ52048
ID AAZ52048 standard; cDNA; 4425 BP.
XX
AC AAZ52048;
XX
DT 18-JUL-2000 (first entry)
XX
DE Rat multidrug resistance protein la cDNA derived from EST sequences.
XX
KW Multidrug resistance protein la; mdrla; multi-specific drug transporter;
KW drug formulation; formulation excipient; compound design; inflammation;
KW pharmacokinetic; oral absorption; cancer; cardiovascular disease;
KW central nervous system disorder; auto-immune disease; kidney disease;
KW EST; expressed sequence tag; ss.
XX
OS Rattus rattus.
XX
FH Key Location/Qualifiers
FT CDS 352..4170
FT /*tag= a
FT /product= "Rat multidrug resistance protein"
XX
PN WO200015650-A1.
XX
PD 23-MAR-2000.
XX
PF 10-SEP-1999; 99WO-US20770.
PR 17-SEP-1998; 98US-0156800.
PR 09-DEC-1998; 98US-0208809.
XX
PA (SMIK) SMITHKLINE BEECHAM CORP.
XX
PI Brun KA, Chenery RJ, Ellens H, Feild JA, Yue L;
XX
DR WPI: 2000-271372/23.
DR P-FSDB; AAY70597.
XX
PT Isolated rat mdrla polynucleotides and polypeptides, useful in assays
PT to provide information on drug formulation, selection of formulation
XX excipients and compound design -
XX
PS Claim 11; Page 27-29; 33pp; English.
XX
CC The present cDNA sequence encodes rat multidrug resistance protein la
CC (mdrla). This cDNA is derived from EST (expressed sequence tag)
CC sequences. Mdrla is used in assays to provide information on
CC drug formulation, selection of formulation excipients and compound
CC design. They are used in cell based, membrane based, binding or other
CC assays to provide information that may enhance drug formulation. This
CC invention further relates to the generation of in vivo and in vitro
CC comparison data to predict oral absorption and pharmacokinetics. This
CC enables the selection of drugs with optimal pharmacokinetics. i.e. good
CC oral bioavailability, brain penetration, plasma half life, and minimum
CC drug interaction. Transgenic and knock-out animals created using rat
CC mdrla provides an insight into treating and preventing human diseases
CC including cancer, inflammation, cardiovascular disease, central nervous
CC system disorders, auto-immune and kidney disease.
SQ Sequence 4425 BP; 1250 A; 992 C; 1113 G; 1067 T; 2 U; 1 other;
Alignment Scores:

Pred. No.: 1.52e-182 Length: 4425
Score: 1882.00 Matches: 362
Percent Similarity: 74.37% Conservative: 137
Best Local Similarity: 53.95% Mismatches: 158
Query Match: 56.64% Indels: 14
DB: 21 Gaps: 5

US-09-873-409-1 (1-659) x AAZ52048 (1-4425)

Qy 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrSer 20
Db 2140 GTCAATGTGGACGAAGAAATCATGATGAGCTCATGAGAGAGAGAAATTTACTTCAA 2199
Qy 21 LeuValMetSerGln-----AspIleLysLysAlaAspGluGlnMetGluSer 36
Db 2200 CTTGTGCATGACTCAGACAGCAGCAAGAAATGAATTAAGAAATGAAGCTTGTGAATCT 2259
Qy 37 -----MetThrTyrSerThrGluArgLysThrAsnSerLeu----- 48
Db 2260 AAAGAUGAAATGATAATGTGGACATGCTCTCAAAAGAAATTCRGGATCCAGTCTAATAAGA 2319
Qy 49 ProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThr 68
Db 2320 AGAAGATCAACTCGCAAAAGCATCGTGGGCCACATGATCAAGACGGGAACTTAGCACC 2379
Qy 69 ---GlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 87
Db 2380 AAAGAGGCTCTGGATGACGACGATCACTCCAGCTTCTTTGGCGGATCTGAAAGTTGAAT 2439
Qy 88 LysProGluTyrProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107
Db 2440 TCAACTGAATGGCCCTTTATTTGGTGGTGTGATTTTGTGGCCATAATTAATGAGGCTTG 2499
Qy 108 HisProValPheSerIleLeuPheAlaLysIleIleThrMetPheGlyAsnAsnAspLys 127
Db 2500 CAACAGCATCTTCCATAATATTTCAAGAGTGTAGGGTTTTTACAAAATGACACC 2559
Qy 128 ThrThrLeu---LysHisAspAlaGluLysIleThrSerMetIlePheValIleLeuGlyVal 146
Db 2560 CCTGAATCCAGCGGAGCAAGCAACTGTGTTCTTTATTTCTGATCTCTGGGATC 2619
Qy 147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 166
Db 2620 APTCTTTTCATTACGCTTTTCTCCAGGCTTTCACATTTGGCAAGCTGGAGATCTCTC 2679
Qy 167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 186
Db 2680 ACCAAGCACTCCGATACATGCTCTTCAATCCATGCTGAGACAGACATAGCTGGTTT 2739
Qy 187 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206
Db 2740 GATGACCTAAAAACACACAGGAGGCTGACCCAGGCTTGGCCAATGACGCTGCTCAA 2799
Qy 207 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226
Db 2800 GTGAAAGGGGCTACAGGGTCTAGGCTTGTGTTATTTACCCAGAACATAGCAAAATCTGG 2859
Qy 227 LeuSerValIleIleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIle 246
Db 2860 ACAGCATCATATATCCCTGATCTACGGCTGGCAATGACACTTTTACTCTAGCAAT 2919
Qy 247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266
Db 2920 GTTCCCATCTATGCTATAGCAGGAGTGGTTGAATGAAATGTAATGTTGTGGACAAGGCTG 2979
Qy 267 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286
Db 2980 AAAGATAAGAGAACTAGAGGTTCTGGGAAGATCGCTTACAGAAAGCAATTTGAAAACTTT 3039
Qy 287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGluMetTyrGluGluMetLeu 306
Db 3040 CGCACTGCTCTCTTTGACTCGGAGCGAAGAGTTTGAATCTATGTATGCCAGAGCTTG 3099

Qy 307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326
Db 3100 CAGATACATACAGAAATGCTTTTGAAGAAAGCGCACGCTTTGGGATCAGCTTTCTCTTC 3159
Qy 327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 346
Db 3160 ACCCAGGCCATGATGATTTCTCTCTATGCTGTGTTCCGTTTGATGCTTACTTGGTG 3219
Qy 347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366
Db 3220 GCACGAGAACTCATGACATTTGAAATGTTCTGTAGTATTCTCAGCTATTGTCTTTGGT 3279
Qy 367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386
Db 3280 GCCATGGCAGTGGGGCAGGTCAGTTTCATTCCTCTGACTACGGAAGCCAAAGCTCTCG 3339
Qy 387 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 406
Db 3340 GCATCCCATCATCAGGATCATTTGAGNAATCCTTGATTTGACAGTACAGCACGAG 3399
Qy 407 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPheTyr 426
Db 3400 GGCCTTGAAGCCTAATATGTTGGAAGAAATGTGAAATTTAATGGAGTCATGTTCAACTAT 3459
Qy 427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446
Db 3460 CCCACCCGACCACTCCAGTCTTCAGGGCTGAGCCTAGAGGTGAAGAGGGGAC 3519
Qy 447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 466
Db 3520 ACCTGGCCCTCTGGGCGAGCAGTGGCTGCGGAAGAGTACAGTGTCCAGTCTGCTTGAG 3579
Qy 467 ArgLeuTyrAspProValGlnGlyValLeuPheAspGlyValAspAlaLysGluLeu 486
Db 3580 CGCTTCTATGACCCATCGCGCAACAGTGTCTAGATGGCAAGAAATAAGCAACTC 3639
Qy 487 AsnValGlnTyrLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506
Db 3640 AATGTCCAGTGGCTCCGCGCCACCTGGGCAATGTGTCCAGGAGCCCATCTGTGTTGAC 3699
Qy 507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValProLeuAspGlu 526
Db 3700 TCGAGCATCGCGAGAACATTCCTACGAGACAAACACCGCTGTCTGTCTCATAGAGAG 3759
Qy 527 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546
Db 3760 ATCGTGAAGGCGACCAAGGAGGCCAACATCCACAGTTTCATCGACTCACTGCTGAGAA 3819
Qy 547 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 566
Db 3820 TACAACACGAGTGGGAGACAAAGGACTCAGCTGTCCGGCGGCGAGAGCAGCGCATC 3879
Qy 567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuAspGluAlaThrSer 586
Db 3880 GCCATCGCGCGCCCTCGTCAGACAGCTCACATCTTACTTCTGATGAAGCGACATCA 3939
Qy 587 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaAargThrGly 606
Db 3940 GCTCTGGATACGAGAGTGAAGAGTCTCCAGGAAGCGCTGGCAAGCCAGCGGAGGC 3999
Qy 607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626
Db 4000 CGCACTGCATTTGATCGCGCAGCGCTGTCCACATCCAGAACCGCAGACTTGTATCGTG 4059
Qy 627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 646
Db 4060 GTGATTCAGAACGGCCAGGTCAAGGAGCACCGCACCCACCGAGCTGCTGCCCCAGAAA 4119
Qy 647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657
Db 4120 GGCATCTATTCTCGATGCTGAGTGCAGGCT 4152

RESULT 8

AAF86127 standard; cDNA; 4186 BP.

XX AAF86127; 100..3942

XX AC

XX DT 25-JUN-2001 (first entry)

XX Cynomologous monkey P-glycoprotein cDNA.

DE DE Cynomologous monkey; P-glycoprotein; PGP; multidrug transporter; MDR1;

XX KW efflux pump; ss.

XX OS Macaca fascicularis.

XX FH Key Location/Qualifiers

XX CDS 100..3942

FT /*tag= a

FT /product= "PGP"

FT /note= "P-glycoprotein"

XX WO200123565-A1.

XX PN 05-APR-2001.

XX PD 28-SEP-2000; 2000WO-US26592.

XX PF 28-SEP-1999; 99US-0156921.

XX PR 12-OCT-1999; 99US-0158818.

XX PA (GENT-) GENTEST CORP.

XX PI Stocker PJ, Steimel-Crespi DT, Crespi CL;

XX WPI; 2001-316136/33.

XX P-PSDB; AAB81064.

XX Novel isolated nucleic acid encoding cynomologous monkey P-glycoprotein (PGP) and homologous PGP polypeptides are useful for predicting bioavailability of compound and increasing PGP transporter activity in cell -

XX Example 1; Page 51-57; 84pp; English.

XX This invention relates to a polynucleotide sequence encoding a cynomologous monkey P-glycoprotein (PGP), and an allelic variant of the PGP protein. PGP, also known as multidrug transporter, MDR1 is a member of the ABC transporter superfamily. The enzyme serves as an efflux pump exporting small molecules across the cell membrane. The invention includes a cynomologous monkey (Macaca fascicularis) PGP coding sequence and protein, and also that of an allelic variant. The PGP polynucleotide sequence is useful for increasing PGP transporter activity in a cell. Antisense sequences of the cDNA are useful for inhibiting PGP transport activity in a mammalian cell. They may also be used for increasing the bioavailability of a drug. The present sequence represents cDNA encoding cynomologous monkey P-glycoprotein.

XX Sequence 4186 BP; 1226 A; 799 C; 1039 G; 1122 T; 0 other;

Alignment Scores:

Pred. No.:	1.17e-191	Length:	4186
Score:	1873.00	Matches:	354
Percent Similarity:	74.29%	Conservative:	146
Best Local Similarity:	52.60%	Mismatches:	155
Query Match:	56.36%	Indels:	18
DB:	22	Gaps:	5

US-09-873-409-1 (1-659) x AAF86127 (1-4186)

Qy 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20

Db 1912 GTCATTGGGAGAAAGGAATCATGATGAGCTCATGAAGAAAGGCAATTTACTTCAA 1971

Qy 21 LeuValMetSerGlnAspIle-----LysLysAlaAspGlu---- 32

Qy 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyTyrSer 20
 :
 Db 1912 GTCATTGGAGAAAGGAATCATGTATGAGCTCATGAAGAAGACGATTACTTCAA 1971

Qy 21 LeuValMetSerGlnAspIle-----LvslvsAlaaaspGlu--- 32

Qy	425	PhcYrPrOCysArgProAspValPheIleLeuAtrGlyLeuSerIeuSerIleGluArg	444
Db	3235	AACATATCCACCCGACGTGGACATATCCAGTGCTTCAGGGGCTGAGCTGGAAGTGAAGAAAG	3299
Qy	445	GlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeu	464
Db	3295	GGCCAGACGCTGGCCCTGGTGGGACAGTGGCTGTGGGAAGACGACGGTGGTCCAGCTC	3355
Qy	465	LeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLys	484
Db	3355	CTGGAGCGGTTCATGACCCCTTTGGCGGGAAAGTGCTCTCTTGACGGCAAGAAATAAAG	3411
Qy	485	GluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnLeuProValLeu	504
Db	3415	CACTGAAATGTTGAGTGGCTCCGAGCACACCTGGGCATCGTCTCCAGGAGGCCATCTGT	3476
Qy	505	PheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeu	524
Db	3475	TTTGACTGCAGCATTAGTGAGAACATTTGCTATGGAGACAAACAGCGGGGTGGTGTACAG	3533
Qy	525	AspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluClyLeuPro	544
Db	3535	GAAGAGATCTGTGAGGGCAGCCAGGAGGGCAATATACACGCCTTCATCAGTTCATCGCT	3599
Qy	545	GluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyClyGlnLysGln	564
Db	3595	AATAAATATAGCACCCAGAGTAGAGACAAAGGNACTCAGCTCTCTGGTGGCCAGAAACAA	3655
Qy	565	ArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuAspGluAla	584
Db	3655	CGCATGCCATAGCTCGTGCCCTTGTATAGACGCTCATATTTTGGCTTTGGATGAAGCC	3711
Qy	585	ThrSerAlaLeuAspAsnAspSerGluLysValValGlnPheAlaLeuAspLysAlaArg	604
Db	3715	ACATCAGCTCTGGATACAGAAAGTGAAAGGTTGTCCAGAGAGCCCTGGACAAAGCCAGA	3776
Qy	605	ThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeu	624
Db	3775	GAAGGCGGTACCTGCATTTGATTTGTCTCCACGCTGTCCACCATCCAGAAATGCAGACTTA	3833
Qy	625	IleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlnLeuLeuArg	644
Db	3835	ATAGTGTGTTTCAGATATGCAGAGTCAGAGAGCACGGCACATCAGCAGCTGTGTGCCA	3899
Qy	645	AsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer	657
Db	3895	CAGAAAGGCATCTATTTTCAATGGTCAAGTGTCCAGGCT	3933

RESULT 10	
AAQ70907	
ID	AAQ70907 standard; DNA; 2726 BP.
XX	
XX	
AC	AAQ70907;
XX	
XX	
DT	27-MAR-1995 (first entry)
XX	
XX	
DE	Multidrug-resistance gene MDR-1.
XX	
XX	
KW	Multidrug-resistance: MDR-1 gene; ds.

27-MAR-1995 (first entry)
Multidrug-resistance gene MDR-1.

DE Multidrug-resistance gene MDR-1.
XX
KW Multidrug-resistance: MDR-1 gene: ds.

XX
OS
SyntheticXX
DN
W00417306-7

XX

XX

XX
XX

EX 27-OCT-1993; 93US-0010433.
PR 16-DEC-1993; 93US-0168621.

XX PA (ONCO-) ONCOR INC.

Db 2057 GCCATAGCTGCGCCCTTGTAGACAGCCTCATATTTTCTTTTGGATGAGCCAGCTCA 2116
 Qy 587 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606
 Db 2117 GCTCTGATACAGAAAGTGAAGAAGTTGTCCAAAGAGCCCTGGACAAAGCCAGAGAGGC 2176
 Qy 607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626
 Db 2177 CGCACCTGCATTGTGATTGCTCACCCGCTGTCCACCATCCAGAATGCAGACTTAATAGTG 2236
 Qy 627 ValLeuHisAsnGlyLysIleLysGlnGlnGlyThrHisGlnGluLeuLeuArgAsnArg 646
 Db 2237 GTGTTTCAGATGGCAGATCAAGGAGCATGGCCACCATCAGCAGCTGTGCGACAGAAA 2296
 Qy 647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657
 Db 2297 GGCATCTATTTTCAATGGTCAGTGTCCAGGCT 2329

RESULT 11

AAQ070916

ID AAQ070916 standard; DNA; 2726 BP.

AC AAQ070916;

DT 27-MAR-1995 (first entry)

DE Multidrug-resistance gene MDR-1.

KW Multidrug-resistance; MDR-1 gene; ds.

OS Synthetic.

PN WO9417210-A.

PD 04-AUG-1994.

PF 25-JAN-1994; 94WO-US00748.

PR 27-JAN-1993; 93US-0010433.

PR 16-DEC-1993; 93US-0168621.

XX (ONCO-) ONCOR INC.

XX Bhatnagar SK, George AL;

XX WPI; 1994-264122/32.

XX Enzymatic amplification of target nucleic acid sequences to form

PT exact or modified copies - has increased fidelity and can identify

PT point mutations or allele(s)

XX Disclosure; Page 30; 69pp; English.

CC This gene is used as an example of a target DNA in a method for the

CC amplification of nucleic acid sequences in a mixture using various

CC DNA probes/primers. The method may be used to detect a mutation or

CC allele in the target, e.g. that that causes a genetic disease.

XX Sequence 2726 BP; 818 A; 538 C; 642 G; 728 T; 0 other;

SQ

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

Length:

Matches:

Conservative:

Mismatches:

Indels:

Gaps:

2726

356

142

159

14

4

(1-2726)

1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20

Db 317 GTCAATTGTGAGAAAGGAATCATGATGAACCTCATGAAAGAGAAAGGCATTACTTCAAA 376

Qy 21 LeuValMetSerGln-----AspIleLysLysAlaAspGluGlnMet 34
 Db 377 CTTGTCAACATGCAGACAGCAGAAATGAAGTTGAATTAGAAATGCAGCTGATGAATCC 436
 Qy 35 GluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--- 53
 Db 437 AAAAGTGAATGTGATGCTTGGAAATGCTTCAAAATGATTCAGATCCAGTCTAATAAGA 496
 Qy 54 -----LysSerIleLys---SerAspPheIleAspLysAlaGluGluSer 67
 Db 497 AAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAGCCCAAGCAGAAAGCTTAGTACC 556
 Qy 68 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 87
 Db 557 AAAAGAGGCTCTGGATCAAAAGTATACCTCCAGTTTCTCTTTGGAGGATTATGAAGCTAAAT 616
 Qy 88 LysProGluTyrProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107
 Db 617 TTAAGTGAATGGCCTTATTTTGTGGTGTATTTGTGCCATTATATAATGGAGCCTG 676
 Qy 108 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAsp 126
 Db 677 CAACCCAGCATTTGCAATAATATTTCAAGATATATAGGGTTTTTACAAGAATTGATGAT 736
 Qy 127 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 146
 Db 737 CCTGAACCAAAACGACAGATAGTAACCTTGTTCACCTATTTCTAGCCCTTGGAAAT 796
 Qy 147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 166
 Db 797 ATTTCTTTTATACATTTTCTTCAGGGTTTCACATTTGGCAAGCTGGAGAGATCCTC 856
 Qy 167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 186
 Db 857 ACCAAGCGGCTCCGATACATGTTTCCGATCCATGCTCAGACAGAGATGTGAGTTGTTT 916
 Qy 187 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206
 Db 917 GATGACCCCTAAAACACCACTGGAGCATTTGACTACAGGCTCGCCAATGATGTGCTCAA 976
 Qy 207 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226
 Db 977 GTTAAGGGGCTATAGTTCAGGCTTCTGTAATATACCAAGATATAGCAATCTTGGG 1036
 Qy 227 LeuSerValIleIleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIle 246
 Db 1037 ACAGGAATAATATATATCTTCTCATCTATGTTGGCAACTAACACTGTACTCTTAGCAAT 1096
 Qy 247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266
 Db 1097 GTACCCCATCATTCGAATAGCAGGAGTTGTTGAAATGAAATGTTGTCTGCAAGCACTG 1156
 Qy 267 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286
 Db 1157 AAAGATAAGAAAGAACTAGAGGTCTGGAAAGATCGCTACTGAAAGCAATAGAAACTTC 1216
 Qy 287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306
 Db 1217 CGAACCGTTGTTTCTTGTGACTCAGGAGCAGAGATTGTAACATATGATGCTCAGAGTTG 1276
 Qy 307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326
 Db 1277 CAGGTACCATACAGAAACTCTTTTGGAGAAAGCACACATCTTTTGGAAATTACATTTTCTTC 1336
 Qy 327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyValaTyrLeuIle 346
 Db 1337 ACCAGGCAATGATGATTTTCTTATGCTGATGTTTCCGGTTTGGAGCCCTACTTGGTG 1396
 Qy 347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366
 Db 1397 GCACATAAACTCATGAGCTTTGAGGATGTTCTGTGTAGTATTTTTCAGCTGTTGTTGGT 1456

QY 367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386
 Db 1457 GCCATGCGCGTGGGCAAGTCAGTTCTCTGCTATGCTATGCAAGCAAAATATCA 1516
 QY 387 AlaAlaHisLeuPheAlaLeuLeuGluLysPheProAsnIleAspSerArgSerGlnGlu 406
 Db 1517 GCAGCCACATCATCATGATCATTTGAAACCAACCCCTTTGATTGACGCTACAGCAGCGAA 1576
 QY 407 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 426
 Db 1577 GGCTAATGCGGAACACATGGAAGGAATGTCATTTGGTGAAGTTGTATTCAACTAT 1636
 QY 427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446
 Db 1637 CCCACCCGACCGACATCCAGTCTTCAGGACTGAGCCTGGAGGTGAAGAGGGCCAG 1696
 QY 447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuGln 466
 Db 1697 ACGTGGCTCTGTGGGCGAGTGGCTGTGGAGAGACACAGTGGTCCAGCTCCTGGAG 1756
 QY 467 ArgLeuTyrPheProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 486
 Db 1757 CGGTTCACGACCCCTTGGCAGGGAAGTGTCTGTGATGCAAGAAATAAACGCACTG 1816
 QY 487 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506
 Db 1817 AATGTTCAAGTGGCTCCGAGCACCTGGGCATCGTGTCCAGGAGCCCATCTGTTGAC 1876
 QY 507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 526
 Db 1877 TGCAGCATGCTGAGAACATTCCTTATGGAGACAAACAGCCGGTGTGTCCACGGAAGAG 1936
 QY 527 IleLysGluAlaAlaAsnAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546
 Db 1937 ATCGTGGGCGAGCAAGAGGCCAACATACATGCTTCATCGAGTCACTGCCTAATAAA 1996
 QY 547 TyrAsnThrGlnValGlyLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 566
 Db 1997 TATAGCACTAAGTAGGAGACAAAGAACTCAGCTCTCTGTGGGCCAGAAACAGGCATT 2056
 QY 567 AlaIleAlaArgAlaLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 586
 Db 2057 GCCATAGCTCGTCCCTTGTAGACAGCTCATATTTGCTTTTGGATGAAGCCACGTCA 2116
 QY 587 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606
 Db 2117 GCTCTGGATACAGAAAGTGAAGAGTTGTCCAAAGAGCCCTGGACAAAGCCAGAGAGGC 2176
 QY 607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspIleVal 626
 Db 2177 GCACCTGCATTTGTGATGCTCACCCTGTCTCCACCATCCAGATCCAGACTTATATAGT 2236
 QY 627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuArgAsnArg 646
 Db 2237 GTGTTTCAGATGGCAGAGTCAAGGAGCATGGCAGCATCGCATCGAGTGTGTGGCAGAAA 2296
 QY 647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657
 Db 2297 GGCATCTATTTTCAATGCTCAGTGTCCAGGCT 2329
 RESULT 12
 AAT43322
 ID AAT43322 standard; DNA; 2726 BP.
 XX
 AC AAT43322;
 XX
 DT 01-SEP-1997 (first entry)
 XX
 DE Multidrug resistance gene-1.
 XX
 KW PCR; primer; amplify; polymerase chain reaction; ligase chain reaction;
 KW LCR; human; multidrug resistance gene; MDR-1; ds.
 XX

OS Synthetic.
 XX WO9639537-A1.
 XX
 PD 12-DEC-1996.
 XX
 PF 04-JUN-1996; 96WO-US08841.
 XX
 PR 05-JUN-1995; 95US-0461823.
 XX
 XX (ONCO-) ONCOR INC.
 XX
 PI Bhatnagar SK, George AL, Nazarenko I;
 XX
 DR WPI; 1997-043158/04.
 XX
 PT Amplification method avoiding strand displacement by polymerase -
 PT used in the detection of mutation(s) and allele(s) associated with
 PT genetic disease and cancer
 XX
 PS Example 1; Page 50-51; 92pp; English.
 XX
 CC This sequence represents the multidrug resistance gene (MDR-1) amplified
 CC by the primers shown in AAT43320 and AAT43321. This sequence can be used
 CC as a target in the method of the invention, for enzymatically amplifying
 CC a target nucleic acid (TNA) sequence contained in a nucleic acid or
 CC mixture of nucleic acids while avoiding strand displacement by
 CC polymerase. The method comprises using three primers, one complementary
 CC to a first segment of the TNA, a second complementary to a second segment
 CC of the TNA, which is adjacent to the first primer, and a third which is
 CC similar to the first segment of the TNA, and is complementary to a
 CC portion of the first primer. The first two primers are hybridised to the
 CC TNA, and a fused amplification product is created from the TNA using
 CC these two primers. The fused product is dissociated, and hybridised to
 CC the third primer, which is then extended. The extended modified
 CC amplification product is hybridised to the first two primers, and these
 CC two primers are ligated. Each of the primers may be labelled using a
 CC different label so that the method can be used to detect the presence of
 CC a mutation or allele by detecting whether the labelled primer is
 CC contained within the fused amplification product or the extended
 CC amplification product. The method of the invention combines certain
 CC aspects of ligase chain reaction (LCR) and polymerase chain reaction
 CC (PCR), but is improved compared to LCR due to the reduced number of
 CC primers needed, and the fact that the entire TNA sequence does not need
 CC to be known.
 XX
 SQ Sequence 2726 BP; 818 A; 538 C; 642 G; 728 T; 0 other;

Alignment Scores:
 Pred. No.: 7.82e-182 Length: 2726
 Score: 1872.00 Matches: 356
 Percent Similarity: 74.22% Conservative: 142
 Best Local Similarity: 53.06% Mismatches: 159
 Query Match: 56.33% Indels: 14
 DB: 18 Gaps: 4
 US-09-873-409-1 (1-659) x AAT43322 (1-2726)

QY 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
 Db 317 GTCATTGTGGAGAAAGAAATCATGATGAACCTCAAGAGAGAAAGGATTTACTTCAA 376
 QY 21 LeuValMetSerGln-----AspIleLysLysAlaAspGluGlnMet 34
 Db 377 CTTGTCAATGCAGACAGCAGGAAATGAAGTTGAATTAAGAAATGCAGCTGATCAATCC 436
 QY 35 GluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--- 53
 Db 437 AAAAGTGAATTTGATGCTTGGAAATGCTTCAATGATTCAGATCCAGTCTAATAAGA 496
 QY 54 -----LysSerIleLys---SerAspPheIleAspLysAlaGluGluSer 67
 Db 497 AAAAGATCAACTCGTAGGAGTGTCCGTGGATCAACAAGCCCAAGACAGAAAGCTTAGTACC 556

XX Kreutzer R, Limmer S, Rost S, Hadwiger P;
 XX WPI; 2002-270454/32.
 XX
 PT Inhibiting gene expression in cells, useful for e.g. treating tumors,
 PT by introducing double-stranded complementary oligonucleotides having unpaired
 PT terminal bases -
 XX
 PS Claim 13; Page 28-30; 104pp; German.
 XX
 CC The invention relates to a method for inhibiting expression of a target
 CC gene (ABL91658-ABL91797) in a cell by introducing at least one
 CC oligonucleotide that has a double-stranded structure consisting of at
 CC most 49 sequential nucleotide pairs, with at least part of one strand
 CC complementary with the target gene and has at least one end a
 CC single-stranded segment of 1-4 nt. The method provides
 CC oligonucleotides for antisense inhibition of gene expression useful
 CC e.g. for treating tumors but the oligonucleotides may also be
 CC directed against genes present in pathogens (e.g. Plasmodium or
 CC viruses/viroids, pathogenic or human, animals or plants) or against
 CC cytokine, Id, developmental or prion genes. The method provides more
 CC effective inhibition of gene expression than use of known
 CC oligonucleotides, probably because the unpaired overhang increases
 CC stability and thus intracellular concentration.
 XX
 SQ Sequence 3840 BP; 1130 A; 738 C; 957 G; 1015 T; 0 other;

Alignment Scores:
 Pred. No.: 1.31e-181 Length: 3840
 Score: 1872.00 Matches: 356
 Percent Similarity: 74.22% Conservative: 142
 Best Local Similarity: 53.06% Mismatches: 159
 Query Match: 56.33% Indels: 14
 DB: 24 Gaps: 4

US-09-873-409-1 (1-659) x ABL91687 (1-3840)

Qy 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTrpSer 20
 Db 1810 GTCAATGTGGAGAAAGAAATCATGATGACTCATGAAAGAGAAAGGCTTACTTCAA 1869
 Qy 21 LeuValMetSerGln-----AspIleLysAlaAspGlnMet 34
 Db 1870 CTTGTCACAATGCAGACAGCAGAAATGAAGTTGAATAGAAATGCAGCTCATGAATCC 1929
 Qy 35 GluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--- 53
 Db 1930 AAAAGTGAAATGTAGTCCCTTGGAAATGTCTTCAATGATTCAGATCCAGTCTAATAAGA 1989
 Qy 54 -----LysSerIleLys---SerAspPheIleAspLysAlaGluSer 67
 Db 1990 AAAAGATCAACTCGTAGGAGTCTCGTGGATCAAGCCCAAGAGCTTAGTACC 2049
 Qy 68 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 87
 Db 2050 AAAGAGGCTCTGGATGAAAGTATACCTCCAGTTCTCTTGGAGGATTATGAAGCTAAAT 2109
 Qy 88 LysProGluTyrProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107
 Db 2110 TPAACGTGAATGGCCCTTATTTTGTGTGTGTTATTTGTCATATATATAATGAGCCCTG 2169
 Qy 108 HisProValPheSerIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAsp 126
 Db 2170 CAACAGCATTTTGCATATATATTTTCAAGAGATATAGGGGTTTTTACAGAAATTGATGAT 2229
 Qy 127 LysThrThrLeuLysHisAspAlaGluLysIleTyrSerMetIlePheValIleLeuGlyVal 146
 Db 2230 CCTGAACAACAAACGACAGATAGTAACTTGTGTTTCTACTATTGTTCTAGCCCTTGAAT 2289
 Qy 147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 166
 Db 2290 ATTTCTTTTATTAACATTTTCTTTCAGGGTTCACATTTTGGCAAGCTGGAGAGATCCTC 2349

Qy 167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 186
 Db 2350 ACCAAGCGGCTCCGATACATGTTTCCGATCCATGCTCAGACAGATGTGAGTTGGTTT 2409
 Qy 187 AspGluLysGluAsnSerThrGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206
 Db 2410 GATGACCCCTAAACACACACCTGGAGCATGTACTACCGGCTCGCCAAATGATGCTCTCAA 2469
 Qy 207 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226
 Db 2470 GTTAAAGGGGCTATAGGTTCCAGGCTTCTGTATATACCAGATATAGCAATCTTGGG 2529
 Qy 227 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle 246
 Db 2530 ACAGGAATAATATATATCTTCTATGTTGGCAACTTAACACTGTTACTCTTAGCAATT 2589
 Qy 247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266
 Db 2590 GTACCCATCATTTGCAATAGCAGGAGTTGTTGAAATGAAATGTTGCTGGACAAAGCATG 2649
 Qy 267 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286
 Db 2650 AAAGATAGAAGAACTAGAGGTGCTGGGAAGATCGTACTGAAGCAATAGAAAACCTTC 2709
 Qy 287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306
 Db 2710 CGAACCGTGTGTTCTTCTGACTCAGGAGCAGAAATGTTGAACATATATGCTCAGAGATTG 2769
 Qy 307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326
 Db 2770 CAGGTACCATACAGAAACTCTTGGAGAAAGCACAATCTTTGGAAATTTACATTTTCCCTC 2829
 Qy 327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyValaTyrLeuIle 346
 Db 2830 ACCCAGGCAATGATGATTTTCTATGCTGAGTGTTCGCTTGGAGCCCTACTTGTGTG 2889
 Qy 347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366
 Db 2890 GCACATAAATCATGAGCTTTGAGGAGTCTCTGTTAGTATTTTTCAGCTGTTGTTCTTGGT 2949
 Qy 367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386
 Db 2950 GCCATGGCCGTGGGGCAAGTCAGTTTCATTTCTCTGATATGCCAAAGCCCAAAATATCA 3009
 Qy 387 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 406
 Db 3010 GGAGCCACATCATCATGATCATTTGAAATAAACCCCTTTGATTGACAGCTACAGCAGGAA 3069
 Qy 407 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 426
 Db 3070 GGCCTAATGTCGAAACACATTTGGAAGAAATGTACATTTTGGTGAAGTTGTATTCAACTAT 3129
 Qy 427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446
 Db 3130 CCCACCCGACCGGACATCCCATGCTTCCAGGACTGAGCCCTGGAGGTGAAGAGGCCGAC 3189
 Qy 447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 466
 Db 3190 ACGTGGCTCTGTGGGACAGTGGCTGTGGGAAGACACAGCTGCTCAGCTCTCTGGAG 3249
 Qy 467 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 486
 Db 3250 CGGTTCTACGACCCCTTGGCAGGAAAGTGTGCTGTGATGCAAGAAATAAAGGACCTG 3309
 Qy 487 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506
 Db 3310 AATGTTCACTGGCTCCGACGACACCTGGGCATCGTGTCCAGAGGCCCATCTCTGTTTAC 3369
 Qy 507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 526
 Db 3370 TGCAGCATTTGCTGAGAAATGTCCTATGGAAGACAAACAGCCGGGTGTTGTACAGGAAGAG 3429

Db 2293 ATTTCTTTTATACATTTTCTCTCAAGGTTTCACATTTGGCAAGCTGGAGATCTCTC 2352
 Qy 167 ThrMetArgLeuAtrGHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaItrpPhe 186
 Db 2353 ACCAAGCGCTCCGATACATGTTTTCGATCCATGCTCAGACAGATGTGATGGTTT 2412
 Qy 187 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206
 Db 2413 GATGACCCCTAAAAACACCACTGGAGCATTTGACTACAGGCTCGCCAAATGATGCTCTCAA 2472
 Qy 207 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226
 Db 2473 GTTAAAGGGCTATAGGTTCCAGGCTTCTGTAATTTACCCAGAAATATAGCAAAATCTTGGG 2532
 Qy 227 LeuSerValIleIleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIle 246
 Db 2533 ACAGGAATTAATATATCTTCACTATGTTGGCAACTAACACTGTTACTCTTAGCAAT 2592
 Qy 247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266
 Db 2593 GTACCACATCTTGCATAGCAGGATGTTGAAATGAAATGTTGCTCTGGACAAGCACTG 2652
 Qy 267 LysAspLysGlnLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286
 Db 2653 AAAGATAAGAAAGAACTAGAGGTGCTGGGAAGATCGCTACTGAAAGCAATAGAAAACCTTC 2712
 Qy 287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306
 Db 2713 CGAACCGTGTCTTCTTGACTCAGGAGCAAGAGTTTGAACATATATGATGCTCAGAGTTTG 2772
 Qy 307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326
 Db 2773 CAGGTACCATACAGAACTCTTTGAGGAAGCACACATCTTTTGGAAATTACATTTTCCITTC 2832
 Qy 327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrIleuile 346
 Db 2833 ACCCAGGCAATGATGATTTTCTCTATGCTGATGTTTCCGGTTTGGAGCCTACTTGGTG 2892
 Qy 347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366
 Db 2893 GCACATAAATCTCAGAGCTTTGAGGATGTTCTGTAGTATTTTCAGCTGTTCTTGTGT 2952
 Qy 367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386
 Db 2953 GCCATGGCGTGGGCAAGTCAGTTCAATTGCTCTGACTATGCCAAGCCAAATATATCA 3012
 Qy 387 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 406
 Db 3013 GCAGCCCATCATCATGATCATTTGAAAAAACCCCTTTGATTGACAGCTTACAGCAGCGAA 3072
 Qy 407 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 426
 Db 3073 GGCCTAAATGCCCAACCATGGAAGGAATGTACATTTGGTGAAGTGTATTTCAACTAT 3132
 Qy 427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446
 Db 3133 CCCACCCGACCGGACATCCAGTGTTCAGGACTGAGCTGGAGGTGAAGAGGGCCAG 3192
 Qy 447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 466
 Db 3193 ACGCTGGCTCTGTGGCAGCAGTGGCTGTGGGAAGAGCACAGTGTCTCCAGCTCTCTGGAG 3252
 Qy 467 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 486
 Db 3253 CGGTCTTACGACCCCTTGCAGGGGAAAGTGTCTGTATGGCAAAATAAAGCACTG 3312
 Qy 487 AsnValGlnTyrLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506
 Db 3313 AATGTCAGTGGCTCCGACACACCTGGGCATCTGTCTCCAGGAGCCCATCTCTTTGAC 3372
 Qy 507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 526

Db 3373 TGCAGCATTTGCTGAGAACATTTGCCCTATGGAGACAACAGCCGGTGGTGTCTACAGGAAGAG 3432
 Qy 527 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546
 Db 3433 ATCGTGGGCGAGCAAGAGGAGGCCAACATACATGCTTTCATCGAGTCACTGCCTAATAAA 3492
 Qy 547 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 566
 Db 3493 TATAGCACTAAAGTAGGAGACAAAGAACTCAGCTCTCTGTGGCCAGAAACAACGCATT 3552
 Qy 567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 586
 Db 3553 GCCATAGCTCTGGCTTGTAGACAGCTCATATTTTGTCTTTGTGATGAGCCACGTCA 3612
 Qy 587 AlaLeuAspAsnAspSerGluLysValGlnHisAlaLeuAspLysAlaArgThrGly 606
 Db 3613 GCTCTGGATACAGAAAGTGAAAGGTTTGTCCAAGAGCCCTGGACAAAGCCAGAGAAGC 3672
 Qy 607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626
 Db 3673 CGCACCTCATTTGATTTGCTCACCGCTGTCCACCATCCAGAAATGCAGACTTAATAGTG 3732
 Qy 627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 646
 Db 3733 GTGTTTCAAGATGGCAGAGTCAAGAGCATGCGACCATCAGCATCAGCTGCTGCGACAGAA 3792
 Qy 647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657
 Db 3793 GGCATCTATTTTCAATGTCAGTGTCCAGGCT 3825
 RESULT 15
 AAZ49333
 ID AAZ49333 standard; cDNA; 3860 BP.
 XX
 AC AAZ49333;
 XX
 DT 14-MAR-2000 (first entry)
 XX
 DE Human G185V mutant multidrug resistance-1 (MDR-1) cDNA.
 KW
 KW Multidrug resistance; MDR-1; P-glycoprotein;
 KW transmembrane efflux pump; haematopoietic stem cell; transduction;
 KW bone marrow transplantation; chemotherapy; radiation therapy; cancer;
 KW gene therapy; gene replacement; genetic defect; thalassaemia;
 KW Gaucher's disease; sickle cell anaemia; leukaemia; ex vivo expansion;
 KW cytokine; mutant; ds.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..3843
 FT /+tag= a
 FT /product= "Human G185V mutant MDR-1 protein"
 XX
 PN WO9961589-A2.
 PD 02-DEC-1999.
 XX
 PF 27-MAY-1999; 99WO-US11825.
 XX
 PR 28-MAY-1998; 98US-0086988.
 XX
 PA (SUJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
 XX
 PI Sorrentino B, Bunting K;
 XX
 DR WPI; 2000-072615/06.
 DR P-PSDB; AAY58187.
 XX
 PT Ex vivo expansion of hematopoietic stem cells transduced with a
 PT sequence encoding human multidrug resistance-1, used for bone marrow
 PT transplantation -

[illegible]


```
Qy 527 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3433 ATCGTAGGGCAGCAAGAGGCCAACATACATGCTTCATCGAGTCACGCTAATAAA 3492

Qy 547 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 566
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3493 TATAGCACTAAAGTAGGAGACAAAGAACTCAGCTCTCTGGTGGCCAGAAACAAGCATT 3552

Qy 567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 586
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3553 GCCATAGCTCGTGCCCTTGTAGACAGCTCATATTTTGCTTTTGATGAAGCCACGTCA 3612

Qy 587 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3613 GCTCTGGATACAGAAAGTGAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGC 3672

Qy 607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3673 CGCACCTGCATTGTGATTGCTCACCGCCTGTCCACCACCAATGCCAGACTTAATAGTG 3732

Qy 627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 646
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3733 GTGTTTCAGATGGCAGAGTCAGAGGAGCATGGCAGCATCAGCAGCTGCTGGCACAGAAA 3792

Qy 647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3793 GGCATCTATTTTCAATGTCAGTGTCAGGCT 3825
```

Search completed: March 30, 2003, 03:05:40
Job time : 735.062 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4 p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 30, 2003, 03:22:44 ; Search time 4532.72 Seconds
(without alignments)
2354.618 Million cell updates/sec

Title: US-09-873-409-1

Perfect score: 3323

Sequence: 1 MIAEKGAHAEIMAKRGLYYS.....QELLNRDIYFKLVNAQSVQ 659

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=xlh
-Q=/cgn2_1/USPFO_spool/US09873409/runat_27032003_115420_19240/app_query.fasta_1.7544
-DB=EST -QFMT=fastcap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOFCU=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human4.0.cdi -LIST=45
-LOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09873409 @CGN 1 1 30544 @runat_27032003_115420_19240 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMMap -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hcc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hcc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_othr:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1008	30.3	943	9	AL520322	AL520322 AL520322
2	834	25.1	998	14	BM904842	BM904842 AGENCOURT
3	780	23.5	2676	11	AK014319	AK014319 Mus muscu
4	770.5	23.2	1019	12	BG248052	BG248052 602359987
5	753	22.7	726	12	BG293345	BG293345 602350738
6	751	22.6	944	12	BF796582	BF796582 602258463
7	743.5	22.4	886	17	BH139685	BH139685 ENTNA47TR
8	740.5	22.3	871	17	AZ682350	AZ682350 ENTKB16TF
9	734.5	22.1	932	17	AZ670821	AZ670821 ENTJUN69TF
10	728.5	21.9	947	17	AZ683753	AZ683753 ENTJUN69TF
11	726.5	21.9	886	17	AZ540627	AZ540627 ENTEQ18TF
12	723.5	21.8	880	17	AZ687805	AZ687805 ENTJUN52TF
13	721.5	21.7	939	14	BQ720763	BQ720763 AGENCOURT
14	720.5	21.7	897	17	AZ541090	AZ541090 ENTDS67TR
15	711.5	21.4	913	17	BH155700	BH155700 ENTROS47TR
16	709.5	21.4	834	17	AZ548312	AZ548312 ENTFCG07TR
17	708.5	21.3	853	17	AZ679807	AZ679807 ENTJUN69TF
18	705.5	21.2	823	17	AZ532602	AZ532602 ENTCS50TR
19	697.5	21.0	861	14	BQ717101	BQ717101 AGENCOURT
20	691	20.8	780	12	BG587938	BG587938 EST489713
21	690.5	20.8	899	17	BH154857	BH154857 ENTRO47TF
22	689.5	20.7	891	17	AZ682250	AZ682250 ENTKT68TR
23	688	20.7	1123	11	AY108485	AY108485 Zea mays
24	683.5	20.6	1341	11	AY108285	AY108285 Zea mays
25	683	20.6	600	13	BM486593	BM486593 pgm2n.pk0
26	682.5	20.5	919	17	AZ690701	AZ690701 ENTJUN69TF
27	678	20.4	926	14	BQ123477	BQ123477 EST609053
28	677.5	20.4	795	12	BF133560	BF133560 601900192
29	669	20.1	547	13	BQ307011	BQ307011 BQ307011
30	667	20.1	937	17	BH150760	BH150760 ENTQ42TR
31	663	20.0	835	13	BQ349604	BQ349604 BQ349604
32	662	19.9	636	10	AV962688	AV962688 AV962688
33	661	19.9	747	13	BQ445774	BQ445774 BQ445774
34	659.5	19.8	628	12	BQ080311	BQ080311 H3052B06-
35	659	19.8	852	17	BH720383	BH720383 BOH2294TF
36	657	19.8	715	14	BQ869512	BQ869512 QGD60909.Y
37	657	19.8	785	10	AV709991	AV709991 AV709991
38	655.5	19.7	787	12	BG584063	BG584063 EST485823
39	653	19.7	609	10	AV986144	AV986144 AV986144
40	651.5	19.6	921	17	AZ687628	AZ687628 ENTJUN69TF
41	651	19.6	746	13	BQ380617	BQ380617 BQ380617
42	650	19.6	729	13	BQ353293	BQ353293 BQ353293
43	647	19.5	1813	11	U66688	U66688 Homo sapien
44	641	19.3	820	13	BQ356540	BQ356540 BQ356540
45	639.5	19.2	842	17	AZ671925	AZ671925 ENTMN93TR

ALIGNMENTS

RESULT 1
AL520322
LOCUS
DEFINITION AL520322 LTI NFL004 NBC2 Homo sapiens CDNA clone CS0DB06YC15 5
prime, mRNA sequence.
ACCESSION AL520322
VERSION AL520322.1 GI:12783815
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 943)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization

JOURNAL
COMMENT

Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: secre@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source

Location/Qualifiers
1..943
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DB006YC15"
/clone_lib="LTI_NFL004_NBC2"
/sex="male"
/tissue type="neuroblastoma cells"
/lab host="DH10B"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@life.com URL : <http://fulllength.invitrogen.com>"

BASE COUNT 253 a 225 c 253 g 211 t 1 others

ORIGIN

Alignment Scores:

Pred. No.: 3.17e-110 Length: 943
Score: 1008.00 Matches: 194
Percent Similarity: 80.40% Conservative: 48
Best Local Similarity: 64.45% Mismatches: 58
Query Match: 30.33% Indels: 1
DB: 9 Gaps: 0

US-09-873-409-1 (1-659) x AL520322 (1-943)

Qy 358 lIeValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAla 377
Db 5 TTATATTTCAGCTGTTGCTTTGGTGCATGGCGGTGGGCAAGTCAGTTTCATTGCT 64
Qy 378 ProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuLysLys 397
Db 65 CTTGACTATGCCAAAGCCAAAATATATCAGCAGCCCATCATCATGATCATTTGAAAAACC 124
Qy 398 ProAsnIleAspSerArgSerGlnGlnGlyLysLysProAspThrCysGluGlyAsnLeu 417
Db 125 CCTTTGATTGACAGCTACAGCAGCGAAGCCCTTAATGCCGAACACATTTGGAAGGAAATGTC 184
Qy 418 GluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGly 437
Db 195 ACATTGGTGAAGTTGATTCAACTATCCACCACCGACATCCAGTGTTCAGGGA 244
Qy 438 LeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGly 457
Db 245 CTGAGCTGGAGGTGAAGAGGCCAGACGCTGCTGTGTGGGAGGAGTGGCTGGG 304
Qy 458 LysSerThrSerValGlnLeuGlnArgLeuTyrAspProValGlnGlnValLeu 477
Db 305 AAGAGCAGACAGTGGTCCAGCTCTGAGGCGGTTCACGACCCCTTGGCAGGGAAGTGTG 364
Qy 478 PheAspGlyValAspAlaLysGluLeuAsnValGlnTTrpLeuArgSerGlnIleAlaIle 497
Db 365 CTTGATGCCAAGAAATAAGGACATGATGTTCAAGTGGCTCCGAGCACACCTGGGCATC 424
Qy 498 ValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAsp 517
Db 425 GTGTCCAGGAGCCCATCTGTTTGTACGTGACGATTCGTCAGACATTCGCTATGCGAGAC 484
Qy 518 AsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAsnIleHis 537
Db 485 AACACCCCGGTGGTGTACAGGAAGAGATCTGTGAGGCGCAGCAAGAGGCGCAACATACAT 544

Qy 538 SerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGln 557
Db 545 GCCTTCATCGAGTCAGTCCCTTAATAAATATATAGCACTAAAGTAGGAGACAAAGGAACCTCAG 604
Qy 558 LeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLys 577
Db 605 CTCTCTGTGTGGCCAGAAACAACGCAATTCGCTATAGTCTGTGCTTGTAGACAGCCTCAT 664
Qy 578 lIeLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGln 597
Db 665 ATTTTGTCTTTGGATGAAGCCACGTCAGCTCTGGATACAGAAAGTGAAGAGTTGTCCAA 724
Qy 598 HisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSer 617
Db 725 GAAGCCCTGGACAAAGCCAGAGAGCGGCACCTGCTGATTGTGCTCACCGCCTGTCC 784
Qy 618 AlaIleGlnAsnAlaAspLeuIleVal-ValLeuHisAsnGlyLysIleLysGluGlnGln 637
Db 785 ACCATCCAGAAATGCAGACTTAATAGTGTGTTTCAGAAATGGCAGAGTCAAGGAGCATGG 844
Qy 637 YThrHisGlnGlnLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSe 657
Db 845 CAGCATCAGCAGCTGCTGGCAGAGAGGCACTATTTTCAATGTGTCAGTGTCCAGGC 904
Qy 657 r 657
Db 905 t 905
RESULT 2
BM904842 998 bp mRNA linear EST 12-MAR-2002
LOCUS AGENCOURT 6699581 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5557655
DEFINITION 5', mRNA sequence.
ACCESSION BM904842
VERSION BM904842.1 GI:19355221
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 998)
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-@email.nih.gov
Tissue Procurement: ATCC/DCTD/DTP
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LHAM12279 row: f column: 24
High quality sequence stop: 738.
Location/Qualifiers
1..998
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5557655"
/clone_lib="NIH_MGC_72"
/tissue type="melanotic melanoma"
/lab host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pCMV-SPORT6; Site: 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2 Kb. Library constructed by Life Technologies."
BASE COUNT 271 a 223 c 248 g 256 t
ORIGIN
Alignment Scores: 3.26e-89 Length: 998
Pred. No.: 834.00 Matches: 164
Score:

Percent Similarity:	74.12%	Conservative:	68
Best Local Similarity:	52.40%	Mismatches:	77
Query Match:	25.10%	Indels:	4
DB:	14	Gaps:	0

US-09-873-409-1 (1-659) x BM904842 (1-998)

Qy	182	Asp	Leu	Ala	Trp	Phe	Asp	Glu	Leu	Glu	Asn	Ser	Thr	Gly	Leu	Thr	Thr	Leu	Ala	201				
Db	61	GAT	GTG	AGT	TGG	TGT	GAT	GAC	CCCT	CAAAAA	CACCACT	GGAGCA	TTC	GTAC	TACCA	GGCT	CGCC	120						
Qy	202	Ile	Asp	Leu	Ala	Gln	Ile	Gln	Gly	Ala	Thr	Gly	Ser	Arg	Ile	Gly	Val	Leu	Thr	Gln	Asn	221		
Db	121	AAT	GAT	GCT	GCT	CAAG	TTC	AAAGGG	GCT	TAT	AGG	TTC	AGG	CTT	GCT	GT	TAAT	TACCC	AGA	T	180			
Qy	222	Ala	Thr	Asn	Met	Gly	Leu	Ser	Val	Ile	Ile	Ser	Phe	Ile	Tyr	Gly	Trp	Glu	Met	Thr	Phe	241		
Db	181	AT	AGCA	AA	TCT	TGG	CAG	CA	GA	TA	AT	TAT	TCT	T	CAT	CT	TAT	TGG	TGG	CACT	TAA	CACTG	240	
Qy	242	Leu	Ile	Leu	Ser	Ile	Ala	Pro	Val	Leu	Ala	Val	Thr	Gly	Met	Ile	Glu	Thr	Ala	Ala	Met	261		
Db	241	TT	ACT	CT	TAG	CA	ATT	GT	TAC	CCAT	CATT	GC	AAT	TAG	CAG	GAG	TGT	TG	AAAT	TG	AAAT	TGTTG	300	
Qy	262	Thr	Gly	Phe	Ala	Asn	Lys	Asp	Lys	Gln	Leu	Leu	Lys	His	Ala	Gly	Lys	Ile	Ala	Thr	Glu	281		
Db	301	TCT	GG	CA	AG	CAC	TG	AA	GA	TAC	GA	AA	CT	T	AG	AG	G	TCT	TGG	AAG	AT	TGTTG	360	
Qy	282	Ala	Leu	Glu	Asn	Ile	Arg	Thr	Ile	Val	Ser	Leu	Thr	Arg	Glu	Lys	Ala	Phe	Glu	Gln	Met	301		
Db	361	GC	AT	AG	AA	AA	ACT	TC	CG	AT	TTC	T	T	GC	T	CAG	GC	AG	AG	AT	T	TGAA	CATATG	420
Qy	302	Tyr	Glu	Glu	Met	Leu	Gln	Thr	Gln	His	Arg	Asn	Thr	Ser	Lys	Lys	Ala	Gln	Ile	Ile	Gly	321		
Db	421	TAT	GCT	CAG	AGT	TTC	GC	AGT	ACC	AAC	AG	AA	CT	TCT	T	TG	AG	AA	AG	CAC	CA	TCTTTGGA	480	
Qy	322	Ser	Cys	Tyr	Ala	Phe	Ser	His	Ala	Phe	Ile	Tyr	Phe	Ala	Tyr	Ala	Ala	Gly	Phe	Arg	Phe	341		
Db	481	ATT	AC	AT	TTT	CTT	C	ACC	AGC	CA	TG	AT	GT	AT	TTT	T	CT	T	AT	G	T	TTCCTGATG	540	
Qy	342	Gly	Ala	Tyr	Leu	Ile	Gln	Ala	Gly	Arg	Met	Thr	Pro	Glu	Gly	Met	Phe	Ile	Val	Phe	Thr	361		
Db	541	GG	AG	CT	ACT	TGG	TGG	GC	CA	TAA	ACT	CAT	GAG	CTT	T	CAG	GAT	G	TCT	G	T	TAGTATT	TTC	600
Qy	362	Ala</																						

RESULT 3
AK014319

LOCUS	AK014319	2676 bp	mRNA	linear	HTC 19-JAN-2000
DEFINITION	Mus musculus 14, 17 days embryo head cDNA, RIKEN full-length enriched library, clone:322401P09:ATP-binding cassette, sub-family B (MDR/TRAP), member 8, full insert sequence.				
ACCESSION	AK014319.1 GI:12852089				
VERSION	HTC; CAP trapper.				
KEYWORDS	Mus musculus (strain:C57BL/6J) 14, 17 days embryo head cDNA to mRNA, clone lib:RIKEN full-length enriched mouse cDNA library clone:322401P09.				
ORGANISM	Mus musculus				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
REFERENCE	1 Carninci,P. and Hayashizaki,Y.				
AUTHORS	Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.				
TITLE	High-efficiency full-length cDNA cloning				
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)				
MEDLINE	92279253				
PUBMED	10349636				
REFERENCE	2				
AUTHORS	Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.				
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)				
MEDLINE	20499374				
PUBMED	20499374				
REFERENCE	3				
AUTHORS	Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kiteunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kita,A. and Hayashizaki,Y.				
TITLE	RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer				
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)				
MEDLINE	20530913				
PUBMED	11076861				
REFERENCE	4				
AUTHORS	Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Iehii,Y., Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S., Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I., Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R., Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T., Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H., Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G., Tomita,M., Quackenbush,J., Schriml,L.M., Staibul,F., Suzuki,R., Tomita,M., Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H., Balzarrelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N., Carninci,P., de Bonaldo,M., Brownstein,M.J., Bult,C., Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D., Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P., Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P., Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H., Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H., Toyooka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L., Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S. and Hayashizaki,Y.				
TITLE	Functional annotation of a full-length mouse cDNA collection				
JOURNAL	Nature 409 (6821), 685-690 (2001)				
MEDLINE	21085560				
PUBMED	11217851				
REFERENCE	5 (bases 1 to 2676)				
AUTHORS	Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A., Arakawa,T., Balzarrelli,R., Bono,H., Brownstein,M., Bult,C., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hiraoka,T., Hori,F., Hume,D., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Oono,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J.,				


```
Db 1200 TTCATTGGGGGCTCCCTTGTGGCTGGACAGCAGCTGAAAGGGGGAGACCTCATGTCTCTTC 1259
Qy 358 lIeValPheThrAlaIleAlaTyRGlyAlaMetAlaIleGlyLysThrLeuValLeuAla 377
Db 1260 CTGTGGTCTCCAGACAGTACAGAGGTCTATGGCC-----AGCCTCTCTGTCTGTCTT 1313
Qy 378 ProGluTyRserLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLys 397
Db 1314 GGTcAGGTGGTACGTGGGCTGAGTCGGGAGCCCGAGTCTTCGAATACATGCGCCTGAGC 1373
Qy 398 ProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeu 417
Db 1374 CCTGTATCCATCCATGACCGGGGGCTACTGTCATCCCAACAAGAGCATTCGTGGTTCATC 1433
Qy 418 GluPheArgGluValSerPhePheTyRProCysArgProAspValPheIleLeuArgGly 437
Db 1434 ACCTTCCAAATATGCCTCTCAGCTACCCCTGCAGACTTGGCTTCAATGTCTCAAGGAC 1493
Qy 438 LeuSerSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGly 457
Db 1494 TTCACCTGAAGTCCCTCTGGCAAGATTGGCTCTTGTGGGCCAGTCTGGGGGAGGA 1553
Qy 458 LysSerThrSerValGlnLeuLeuGlnArgLeuTyRAspProValGlnGlnValLeu 477
Db 1554 AAGACCAcAGTTCCTCTGCTGAACGCTTCTATGACCCCTGAAGCTGGCTCGGTGAGC 1613
Qy 478 PheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGln---IleAla 496
Db 1614 TTGGATGGCGATGACCTCGCAACTCTCAACCCCTCTGGCTCCGGGGCCAGGTATAGGT 1673
Qy 497 lIeValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyRGly 516
Db 1674 TTCATCACCCAGGACCACTCTGTTTGCACACCATCATGGAGATATCCGATTGGG 1733
Qy 517 -----AspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAla 534
Db 1734 AAGCTGGATGCTTCGGAT-----GAAGAGGTGTACACAGCTGCACGAGAAGCC 1781
Qy 535 AsnIleHisSerPheIleGluGlyLeuProGluLysTyRAsnThrGlnValGlyLeuLys 554
Db 1782 AATGCCCAcGAGTTTCATCAGCAGCTTCCCGATGGCTACAGCACTGTGGTGGTGAAGCGG 1841
Qy 555 GlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuGln 574
Db 1842 GCACAACTTGTCTGTGGCCAGAGCAGCGCCTAGCCATCGCAGTGCCTCATCAAG 1901
Qy 575 LysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 594
Db 1902 CAGCCACAGTCTGATCTGGACGAGGCCACCACTGCGCTAGATGCAGATCCGAGAGG 1961
Qy 595 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 614
Db 1962 GTGGTACAGGAGGCGCTGGACCGGCCAGTGTCTGGCCGACCGCTGTGTGTCTATCCCCAC 2021
Qy 615 ArgLeuSerAlaIleGluAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys 634
Db 2022 CGCTTAGTACTGTCCGTGCACCCCACTCCATCATTTGTATGGCAATGGCCAGTCTGT 2081
Qy 635 GlyGlnGlyThrHisGlnGlnLeuLeuArgAsnArgAspIleTyRPhelLysLeuValAsn 654
Db 2082 GAGGTGGGACCCACGAGAACTCTCTTAAAGAGGGCGGCTCTATTTCAGAGCTTATCCGG 2141
Qy 655 AlaGlnSerVal 658
Db 2142 AGACAAACCCGT 2153
RESULT 4
BG248052 1019 bp mRNA linear EST 13-FEB-2001
LOCUS 60235987F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:4488404 5',
DEFINITION mRNA sequence.
ACCESSION BG248052
```

```
VERSION BG248052.1 GI:12757867
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1019)
NTH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-f@mail.nih.gov
Tissue procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
National Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10334 row: n column: 21
High quality sequence stop: 650.
FEATURES
source
1..1019
Location/Qualifiers
/organism="Mus musculus"
/strain="FVB/N"
/db xref="taxon:10090"
/clone="IMAGE:4488404"
/clone_lib="NCI_CGAP Mam1"
/tissue_type="tumor, biopsy sample"
/dev_stage="10 months, virgin"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"
BASE COUNT 276 a 262 c 283 g 196 t
ORIGIN
Alignment Scores:
Pred. No.: 1,52e-81 Length: 1019
Score: 770.50 Matches: 154
Percent Similarity: 79.17% Conservative: 36
Best Local Similarity: 64.17% Mismatches: 47
Query Match: 23.19% Indels: 3
DB: 12 Gaps: 1
US-09-873-409-1 (1-659) x BG248052 (1-1019)
Qy 414 GluGlyAsnLeuGluPheArgGluValSerPhePheTyRProCysArgProAspValPhe 433
Db 4 GAAGGAATGTGCAATTTAGTGGAGTCGTGTCAACTATCCACCCGACCCAGCATCCCA 63
Qy 434 lIeLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySer 453
Db 64 GTGCTTCAGGGGCTGAGCCTTGAGGTGAAGAGGGCCAGACGCTGGCCTCTGTGGGAGC 123
Qy 454 SerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyRAspProValGln 473
Db 124 AGTGCTCGGGAGAGACACAGTGTGTCCAGCTGTCTCGAGCGCTTCTACGACCCCATGGT 183
Qy 474 GlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSer 493
Db 184 GGATCAGTGTCTTAGATGGCAAGAAATAAGCAACTGAATGTCCAGTGGCTCCGAGCA 243
Qy 494 GlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIle 513
Db 244 CAGCTGGGCATTGTGTCCCAAGAGGCCATTCTCTTTGACTGCAGCATCGCAGAGCAATT 303
Qy 514 AlaTyRGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAla 533
Db 304 GCCTACGAGAGACAAACGCGGGTCTGTCTTATAGGAGATTTGAGGGCCACCCAGGAG 363
Qy 534 AlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyRAsnThrGlnValGlyLeu 553
```


High quality sequence start: 9
 High quality sequence stop: 669.
 FEATURES
 source
 1. 944
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4341710"
 /clone_lib="NIH MGC 85"
 /tissue_type="lymphoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /notes="Organ: lymph; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 1.867 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH MGC Library."
 BASE COUNT 265 a 241 c 237 g 201 t
 ORIGIN

Alignment Scores:
 Pred. No.: 3e-79 Length: 944
 Score: 751.00 Matches: 166
 Percent Similarity: 74.62% Conservative: 31
 Best Local Similarity: 62.88% Mismatches: 60
 Query Match: 12.60% Indels: 9
 DB: 12 Gaps: 2

US-09-873-409-1 (1-659) x BF796582 (1-944)

Qy 380 TyrSerIysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysProAsn 399
 Db 39 TATGCTAAAGCTAAGCTCTCTGCGACCCACCTATTATCATGCTGTTTGAAGACACACCTCTG 98
 Qy 400 IleAspSerArgSerGlnGluGlyLysProAspThrCysGluGlyAsnLeuPhe 419
 Db 99 ATTGACACTACATGACAGGGGCTGAAGCTGTATTAATTGAAGGAAATATACATT 158
 Qy 420 ArgGluValSerPhePheTyProCysArgProAspValPheIleLeuArgGlyLeuSer 439
 Db 159 AATGAAGTCGTGTTCAACTATATCCACCGAGCAACGTCGCCAGTCTTTCAGGGGTGAGC 218
 Qy 440 LeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSer 459
 Db 219 CTGGAGGTGAAGAAAGGCGACACATAGCCCTGGTGGGCGAGCGTGTGGGAAGAGC 278
 Qy 460 ThrSerValGlnLeuLeuGlnArgLeuTyRaspProValGlnGlnValLeuPheAsp 479
 Db 279 ACGTGGTCCAGCTCTGAGGGGTTCACGACCCCTTGGCGGGGACAGTGTCTTCGAT 338
 Qy 480 GlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValPro 499
 Db 339 GGTCAAGAGCAAGAAACTCAATCTCCAGTGGCTCAGAGCTCAACTCGGAATCGTGCT 398
 Qy 500 GlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyRGLysAspAsnSer 519
 Db 399 CAGGAGCCTATCTTATTCAGTGCAGCATTCGCCGAGATATTGCTTCATGAGACACAGC 458
 Qy 520 ArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPhe 539
 Db 459 CGGGTTGTATCAGAGTAAATCGTAGTCGAGCAAGCTGCCAACATACATCTCTTC 518
 Qy 540 IleGluGlyLeuProGluLysTyRAsnThrGlnValGlyLysGlyValAlaGlnLeuSer 559
 Db 519 ATCGAGAGCTTACCCACAAATATGAACAGAGTGGGAGATAGGGGACTCAGCTCTCA 578
 Qy 560 GlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLys-ProLysIleLe 579
 Db 579 GGAGGTCAAAACAGAGGATTCCTATTGCGGAGCCCTCATCAGACACACCTCAATCTC 638
 Qy 579 uLeuLeuAspGluAlaThrSerAla-LeuAspAsnAspSerGluLysValVal-----G 597
 Db 639 CCTGTTGATGAAGCTACATAGCTCTCGATACCTCGACAAGTGAACACGGCTGTGCCAA 698
 Qy 597 InHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuS 617

Db 699 GAAGCCCTGTACACAGCAGAGAGCGCCGACCTGCTGATGCTCACCGCTGTC 758
 Qy 617 erAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnG 637
 Db 759 AAC--ATCCAGATGCAGACTT-ATAGTGGTGTTCCTCCACAGCGGAGAGTCAAG----- 808
 Qy 637 lyThrHis 639
 Db 809 --ACACAT 814
 RESULT 7
 BH139685
 LOCUS ENTNA477R Entamoeba histolytica Sheared DNA Entamoeba histolytica
 DEFINITION genomic, DNA sequence.
 ACCESSION BH139685
 VERSION BH139685.1 GI:15098746
 KEYWORDS GSS.
 ORGANISM Entamoeba histolytica.
 SOURCE Entamoeba histolytica
 ORGANISM Eukaryota; Entamoebidae; Entamoeba.
 REFERENCE 1 (bases 1 to 886)
 Loftus, B., Wang, Z., Van Aken, S. and Fraser, C.
 Determination of clone end sequences from Entamoeba histolytica
 HMI:IMSS sheared DNA library (2001)
 JOURNAL Unpublished (2001)
 COMMENT Contact: Brendan J Loftus
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0208
 Fax: 301 838 3543
 Email: bjlloftus@tigr.org
 Clones are derived from the Entamoeba histolytica HMI:IMSS sheared
 DNA library
 Seq primer: M13-Reverse
 Class: shotgun
 High quality sequence start: 18
 High quality sequence stop: 851.
 FEATURES
 source
 Location/Qualifiers
 1. .886
 /organism="Entamoeba histolytica"
 /strain="HMI:IMSS"
 /db_xref="taxon:5759"
 /clone_lib="Entamoeba histolytica Sheared DNA"
 /note="Vector: pHOsi; Site 1: Bst I; Constructed at The
 Institute for Genomic Research (TIGR), Rockville, MD.
 Genomic DNA isolated from broth cultures of E. histolytica
 using a method described by Clark and Diamond (Clark,
 C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
 method for isolate identification. Exp. Parasitol.
 77:450.). The DNA was mechanically sheared to give a
 tight size distribution (~2 kb). The v + i method used for
 the library construction is described in detail in Smith,
 H.O. and Venter, J.C. (Making small insert libraries for
 whole genome shotgun sequencing projects. In Genome
 Sequencing: A Practical Approach, eds. M. Vaudin and B.
 Borell, Oxford University Press, 1999)."
 BASE COUNT 348 a 105 c 176 g 257 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 2.16e-78 Length: 886
 Score: 743.50 Matches: 151
 Percent Similarity: 73.43% Conservative: 48
 Best Local Similarity: 55.72% Mismatches: 71
 Query Match: 22.37% Indels: 2
 DB: 17 Gaps: 1
 US-09-873-409-1 (1-659) x BH139685 (1-886)
 Qy 387 AlaAlaHisLeuPheAlaLeuGluLysProAsnIleAspSerArgSerGlnGlu 406


```

Db 20 GCAAGATATCTATGATGTTATGATAGAAACCACT-ATTGATTTGTTATAGTGAGAA 78
Qy 407 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 426
Db 79 GGTCAAAACATTAAATGATGTTAAAGGTGAATGAAATTAAGACATTTGTTTCAGATAT 138
Qy 427 ProCysArgProAspValPheLeuArgGlyLeuSerLeuSerLeuSerLeuGluArgGlyLys 446
Db 139 CCAACAGACACACACATCTGCTTGAAGGTATTTTCATTCAAAGTGAACCAAGGAAA 198
Qy 447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuGln 466
Db 199 ACTGTTCATAGTAGGAGCATCAGGATGGTAAATCAACATCAGTTCAAGTTGATTGAA 258
Qy 467 ArgLeuTyrAspProValGlnGlnValLeuPheAspGlyValAlaLysGluLeu 486
Db 259 AGATTTTATGATCCACACATGAGATGTTATTTAGATGACATAATCAAGATTTG 318
Qy 487 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506
Db 319 AATATTCATTCTTAAAGTCAAAATGGATGGTAGGACAAAGACGATTTATTTGCT 378
Qy 507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgVal---ValProLeuAsp 525
Db 379 GAAAGTGTATGATATATATAGAGGAGTACCTAAAGAGTTGAGTAAGTAATGAA 438
Qy 526 GluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGlu 545
Db 439 CAAATTTATGCTGCTGCTAAATGGCAATGTCACATGACTTTTATTTTTCAGCAATGCCAGAA 498
Qy 546 LysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArg 565
Db 499 GGATATAACACATGATGGTGTAGTGTGAGGTGCACAAATTTTCAGGAGCACAAACAAAGA 558
Qy 566 LeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThr 585
Db 559 ATGCTATTGTCAGTGCATGATGATGAATCCAAAGTGTATTATCTCGATGAAGCTACA 618
Qy 586 SerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThr 605
Db 619 TCAGCACTTGATTCAGAAATGAAAGATTTGTACAAAGATGCACTTCACAAAGCAGCAAAA 678
Qy 606 GlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaLeuIle 625
Db 679 GGAAGAACCAACATTTGTAATTCACATAGATTATCAACTATTTCAAATGCAGATCAATA 738
Qy 626 ValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlnLeuLeuArgAsn 645
Db 739 TGTGTTATTATGAGAGAGAAATTCAGAAAGAGGAAACATCAAGAGTTATTAGATTG 798
Qy 646 ArgAspIleTyrPheLysLeuValAsnAlaGln 656
Db 799 AAAGGATTTTATTATACACTTGCTGTCAACAA 831

```

```

RESULT 8
LOCUS AZ682350 871 bp DNA linear GSS 14-DEC-2000
DEFINITION ENTKE16TF Entamoeba histolytica Sheared DNA Entamoeba histolytica
GENOMIC, DNA sequence.
ACCESSION AZ682350
VERSION AZ682350.1 GI:11819496
KEYWORDS GSS.
SOURCE Entamoeba histolytica.
ORGANISM Entamoeba histolytica
REFERENCE 1 (bases 1 to 871)
AUTHORS Loftus,B., Van Aken,S. and Fraser,C.
TITLE Determination of clone end sequences from Entamoeba histolytica
JOURNAL HMI:IMSS sheared DNA library
COMMENT Unpublished (2000)
Contact: Brendan J Loftus
Department of Eukaryotic Genomics

```

The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0208
 Fax: 301 838 3543
 Email: bjoftus@igr.org
 Clones are derived from the Entamoeba histolytica HMI:IMSS sheared DNA library
 Seq primer: M13-Forward
 Class: shotgun
 High quality sequence start: 16
 High quality sequence stop: 860.
 Location/Qualifiers
 1. .871
 /organism="Entamoeba histolytica"
 /strain="HMI:IMSS"
 /db_xref="taxon:5759"
 /clone_lib="Entamoeba histolytica Sheared DNA"
 /note="Vector: PHOS1; Site_1: Bst I; Constructed at The Institute for Genomic Research (TIGR), Rockville, MD. Genomic DNA isolated from broth cultures of E. histolytica using a method described by Clark and Diamond (Clark, C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a method for isolate identification. Exp. Parasitol. 77:450.). The DNA was mechanically sheared to give a tight size distribution (~2 kb). The v + i method used for the library construction is described in detail in Smith, H.O. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds M. Vaudin and B. Barrell, Oxford University Press, 1999)."

BASE COUNT 349 a 117 c 174 g 231 t
 ORIGIN

Alignment Scores:

Pred. No.: 4,83e-78 Length: 871
 Score: 740.50 Matches: 147
 Percent Similarity: 70.04% Conservative: 47
 Best Local Similarity: 53.07% Mismatches: 82
 Query Match: 22.28% Indels: 1
 DB: 17 Gaps: 1

US-09-873-409-1 (1-659) x AZ682350 (1-871)

```

Qy 381 SerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIle 400
Db 39 AATAGTCTAAAGTTGCCGCAATTTAATGTTTATCAACAATGATGATGATCCAGATAT 98
Qy 401 AspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArg 420
Db 99 GATTGTCAGTCTATTGGAGGTGAATGTCCTCACTGATGTAATGGAATATTAGATTGAG 158
Qy 421 GluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeu 440
Db 159 GATGTTCAATTTGTTTATCCACAAGACTGCTCATCATGATTATTAAGGACTTCACCTT 218
Qy 441 SerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThr 460
Db 219 GAAATTAAGAAAGGACAAACAATTCATAGTTAGTGGAGCATCAGGATGTGGGAAGTCACT 278
Qy 461 SerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlnValLeuPheAspGly 480
Db 279 ACTATTCAATTAATCCAAAGAAATTTATGATCCAAATGTTGGAAGAGTAACATTAGACGA 338
Qy 481 ValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGln 500
Db 339 AAAGATATACGAGAGTGAATATCAATGTTTAAAGATCAATAGGATTTAGTTGACAA 398
Qy 501 GluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArg 520
Db 399 GAACCAAGTGTGTTGAGGAGCAAAATTCGAGAAATAATTTATGCTTGGAGCTAAAGAAGA 458
Qy 521 ValValProLeu---AspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPhe 539

```



```
Db 459 GCAACACCAAGTGAAGAGAGATGATTGAATGTGCTAAATGGCAATGCACATGACTTC 518
Qy 540 IleGluGlyLeuProGluLysThrGlnValGlyLeuLysGlyValAlaGlnLeuSer 559
Db 519 ATTTCTAACTTCAGAGGATATGACACATAATTGGAGAAAGGAGCATTTATATCA 578
Qy 560 GlyGlyGlnLysGlnArgLeuAlaAlaArgAlaLeuLeuGlnLysProLysIleLeu 579
Db 579 GGAGGACAAAACAAAGAAATTCGAATTCACGCTGCATGATTGGAACCCCATCTATTCTT 638
Qy 580 LeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAla 599
Db 639 CTTCCTGATGAAGCTACATCAGCCTTGATACACAAGTGAAGAAAGATGTACAAGACCA 698
Qy 600 LeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIle 619
Db 699 CTTGAAAAGCATCTAAAGGAGAACACAAATATTGTAGCACATAGACTAACCACTGTT 758
Qy 620 GlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHis 639
Db 759 AGAAATGCAGATAAAATATGTGTATTCATCAAGGAGAAATATTATTGAACAAGGAACAT 818
Qy 640 GlnGluLeuLeuArgAsnArgAspIleTyPhelysLeuValAsnAlaGln 656
Db 819 CAAGAATTAATGGATTTCGAAGGAACATATTATTATGGATTAGTCAAAAGACAA 869

RESULT 9
AZ670821/c 932 bp DNA linear GSS 14-DEC-2000
LOCUS ENTJUN69TF Entamoeba histolytica Sheared DNA Entamoeba histolytica
genomic, DNA sequence.
ACCESSION AZ670821
VERSION 1.1
KEYWORDS GSS.
SOURCE Entamoeba histolytica.
ORGANISM Entamoeba histolytica.
REFERENCE 1 (bases 1 to 932)
AUTHORS Loftus, B., Van Aken, S. and Fraser, C.
TITLE Determination of clone end sequences from Entamoeba histolytica
HMI:IMSS sheared DNA library
JOURNAL Unpublished (2000)
COMMENT Contact: Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 3543
Email: b.loftus@igr.org
Clones are derived from the Entamoeba histolytica HMI:IMSS sheared
DNA library
Seq primer: M13-Forward
Class: shotgun
High quality sequence start: 46
High quality sequence stop: 867.
FEATURES
source
1..932
/organism="Entamoeba histolytica"
/strain="HMI:IMSS"
/db_xref="taxon:5759"
/clone_lib="Entamoeba histolytica Sheared DNA"
/notes="Vector: pHOSt1; Site 1: Bet 1; Constructed at The
Institute for Genomic Research (TIGR), Rockville, MD.
Genomic DNA isolated from broth cultures of E. histolytica
using a method described by Clark and Diamond (Clark,
C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
method for isolate identification. Exp. Parasitol.
77:450.). The DNA was mechanically sheared to give a
tight size distribution (~2 kb). The v + i method used for
the library construction is described in detail in Smith,
H.O. and Venter, J.C. (Making small insert libraries for
whole genome shotgun sequencing projects. In Genome
Sequencing: A Practical Approach, eds. M. Vaudin and B.
```

```
BASE COUNT 258 a 182 c 125 g 367 t
ORIGIN
Alignment Scores: 2,86e-77 Length: 932
Pred. No.: 734,50 Matches: 150
Score: 71.11% Conservative: 42
Percent Similarity: 55.56% Mismatches: 77
Best Local Similarity: 22.10% Indels: 2
Query Match: 17 Gaps: 1
DB: 1

US-09-873-409-1 (1-659) x AZ670821 (1-932)
Qy 380 TyrSerLysAlaLysSerGlyValAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsn 399
Db 835 TTTGCTACAGCTAAAGCTTCTGCATATATAGAATTTATCAACAATTTGATAGAATCCAGAT 776
Qy 400 IleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPhe 419
Db 775 ATTGATTGTAGATCTACAGCTGGTGAATGTCCAACTGAGTGTAAATGGAAATATTATACATTA 716
Qy 420 ArgGluValSerPhePheTyProCysArgProAspValPheIleLeuArgGlyLeuSer 439
Db 715 GAAGATGTTCAATTTAGATATCCAAAGACCAACTAAACAATTTCTTGGTGACTTGAT 656
Qy 440 LeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSer 459
Db 655 CTTGAATTAAGAAGGACAAACAGTTGCTATTAGTAGGAGCATCAGGATGTGTTAAATCA 596
Qy 460 ThrSerValGlnLeuLeuGlnArgLeuTyArgProValGlnGlyGlnValLeuPheAsp 479
Db 595 ACTACTATTCAATTAGTCCAAAGAAATTTATGATCCAGTTGGTGGTCAAGTCAAAATTTAGAT 536
Qy 480 GlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValPro 499
Db 535 GGAAAGATTTAAGAGATTTAAATATCAATGGTTAAGAAATCAAAATAGGATTTAGTGA 476
Qy 500 GlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyArgLysAsnSer 519
Db 475 CAGAACCTATTGTTGTCATGTACTATTAGAGAATAATATTATCTCTGGAGCTAGAGAT 416
Qy 520 ArgValValProLeu---AspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSer 538
Db 415 GGAGAAACACCAACTCAAGAAAGAGATGTTGAAATGTGCTAAATGGCAATGCACATGAA 356
Qy 539 PheIleGluGlyLeuProGluLysTyArgThrGlnValGlyLeuLysGlyValAlaGlnLeu 558
Db 355 TTTATTTCTCATCTCTCCAGAGGATATGATACAAATGGTAGGAGAAAAGGAGCTGCATTA 296
Qy 559 SerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIle 578
Db 295 TCAGAGGACAAAACAAAGAAATTTGCTATTGCACGTGCTATTGATTAGAAACCTACAAT 236
Qy 579 LeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHis 598
Db 235 TTATTACTTGATGAAGCTACATCAGCACTTGATACACAAGTGAAGAAAGATTTGATCAACAA 176
Qy 599 AlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAla 618
Db 175 GCATTTGAAAAGCTAGTCAAGGAGAAACAACAATTTGTTGAGCACATAGATTAACTACT 116
Qy 619 IleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThr 638
Db 115 GTTGAATATGAAGTAGAATTTGTTGTTTCCATCAAGGAGAAATATTAT-GAACAGGAACA 57
Qy 639 HisGlnGluLeuLeuArgAsnArgAspIle 648
Db 56 CATCAAGAATTAAGGATTTAAAGAGGATTTA 27

RESULT 10
AZ683753
LOCUS AZ683753 947 bp DNA linear GSS 14-DEC-2000
```

Barell, Oxford University Press, 1999).


```

DEFINITION   ENT1196TF Entamoeba histolytica Sheared DNA Entamoeba histolytica
ACCESSION    AZ683753
VERSION      AZ683753.1 GI:11820899
KEYWORDS     GSS.
SOURCE       Entamoeba histolytica.
ORGANISM     Entamoeba histolytica.
REFERENCE    1 (bases 1 to 947)
AUTHORS      Loftus,B., Van Aken,S. and Fraser,C.
TITLE        Determination of clone end sequences from Entamoeba histolytica
JOURNAL      HMI:IMSS sheared DNA library
COMMENT      Unpublished (2000)
              Contact: Brendan J Loftus
              Department of Eukaryotic Genomics
              The Institute for Genomic Research
              9712 Medical Center Dr., Rockville, MD 20850, USA
              Tel: 301 838 0208
              Fax: 301 838 3543
              Email: bjlloftus@tigr.org
              Clones are derived from the Entamoeba histolytica HMI:IMSS sheared
              DNA library
              Seq primer: M13-Forward
              Class: shotgun
              High quality sequence start: 40
              High quality sequence stop: 778.
              Location/Qualifiers
                1..947
                  /organism="Entamoeba histolytica"
                  /strains="HMI:IMSS"
                  /db_xref="taxon:5759"
                  /clone_lib="Entamoeba histolytica Sheared DNA"
                  /note="Vector: pHOS1; Site 1: Bst I; Constructed at The
                  Institute for Genomic Research (TIGR), Rockville, MD.
                  Genomic DNA isolated from broth cultures of E. histolytica
                  using a method described by Clark and Diamond (Clark,
                  C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
                  method for isolate identification. Exp. Parasitol.
                  77:450.). The DNA was mechanically sheared to give a
                  tight size distribution (~2 kb). The v + i method used for
                  the library construction is described in detail in Smith,
                  H.O. and Venter, J.C. (Making small insert libraries for
                  whole genome shotgun sequencing projects. In Genome
                  Sequencing: A Practical Approach, eds. M. Vaudin and B.
                  Barrel, Oxford University Press, 1999)."
BASE COUNT   379 a 121 c 178 g 269 t
ORIGIN
Alignment Scores:
Pred. No.:    1.55e-76      Length:      947
Score:        728.50       Matches:    149
Percent Similarity: 69.82%   Conservative: 43
Best Local Similarity: 54.18% Mismatches:    81
Query Match:  21.92%       Indels:      2
DB:           17          Gaps:         1

US-09-873-409-1 (1-659) x AZ683753 (1-947)

QY  383 AlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleaspSer 402
Db  69 GCAAAGCATGACGCTATATAATATTACAAATGATGATAGATCCAGATATTGATGT 128
QY  403 ArgSerGlnGluLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluVal 422
Db  129 CAGTCTATTGGAGGTGAATGTCCTCACTGAGTGTAAATGGGAATATTAGATTGATGTT 188
QY  423 SerPhePheTyrProCysArgProaspValPheLeuArgGlyLeuSerLeuSerIle 442
Db  189 CAATTGTGTTATCCAAACAGACCATCTCATCATGTATTAAAAGGAATTGACCTTGAAT 248
QY  443 GluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlySerThrSerVal 462
Db  249 AGAAGAGGAAATATTCATTAGTTGGAGCATCAGGATGAGAAAGTCAACTACTATT 308

```

```

QY  463 GlnLeuLeuGlnArgLeuTyrAspProValGlnGlnGlyGlnValLeuPheaspGlyValasp 482
Db  309 CAATTAAATCAAAGAATAATTATGAACCAATGTTGGAGAGTAACATTAGATGGGAAGAC 368
QY  483 AlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluPro 502
Db  369 ATAAGAGAGCTTGAATATCAATGTTTAAAGAAATCAATAGGAATAGTTGGACAAGAACCA 428
QY  503 ValLeuPheAsnCySerIleAlaGluAsnIleAlaTyrGly--AspAsnSerArgVal 521
Db  429 GTATTATTTTCAGGACAATTTGGAGAAATATATCATGCTTGGAGCTTAAAGAGAGAAACA 488
QY  522 ValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGlu 541
Db  489 CTAAGTAAAGAGAGAGATGATTGAATGTACTAAATGGCAATGCACATGACTTTATTTCT 548
QY  542 GlyLeuProGluLysTyrAsnThrGlnValGlyLysGlyAlaGlnLeuSerGlyGly 561
Db  549 AAACCTCCAGAGGATATGACACAATAATTGGAGAAAAAGGAGCATTTATTATCAGGAGGA 608
QY  562 GlnLysGlnArgLeuAlaIleAlaArgAlaLeuGlnLysProLysIleLeuLeuLeu 581
Db  609 CAARAACAAGAATTGCAATTGCATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 668
QY  582 AspGluAlaThrSerAlaLeuAspAsnAspSerGluLeuValValThrHisArgLeuSerAlaIleGlnAsn 601
Db  669 GATGAAGCTTACATCAGCACTTCATACACAAGTGAAGAAATGTTACAGAAGCACTTGAC 728
QY  602 LysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsn 621
Db  729 AAAGCATCAAAAGGAGGAGCAACAATTTATTGTAGGACATAGACTATCGACTATTTCAAAAT 788
QY  622 AlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlu 641
Db  789 GCAGATCAAAATATGTTATTATGAGAGGAAAAAATAGTAGAACAGGAACACATCAAGAA 848
QY  642 -LeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAla 655
Db  849 TTTATTGGATTTGAAAAGATTTTATTATATACACTTGCCTTATGCA 891

RESULT 11
AZ540627
LOCUS      ENTEQ18TF Entamoeba histolytica Sheared DNA Entamoeba histolytica
DEFINITION genomic, DNA sequence.
ACCESSION  AZ540627
VERSION    AZ540627.1 GI:11147603
KEYWORDS   GSS.
SOURCE     Entamoeba histolytica.
ORGANISM   Entamoeba histolytica.
REFERENCE  1 (bases 1 to 886)
AUTHORS    Loftus,B., Van Aken,S. and Fraser,C.
TITLE      Determination of clone end sequences from Entamoeba histolytica
JOURNAL    HMI:IMSS sheared DNA library
COMMENT    Unpublished (2000)
              Contact: Brendan J Loftus
              Department of Eukaryotic Genomics
              The Institute for Genomic Research
              9712 Medical Center Dr., Rockville, MD 20850, USA
              Tel: 301 838 0208
              Fax: 301 838 3543
              Email: bjlloftus@tigr.org
              Clones are derived from the Entamoeba histolytica HMI:IMSS sheared
              DNA library
              Seq primer: M13-Forward
              Class: shotgun
              High quality sequence start: 23
              High quality sequence stop: 886.
              Location/Qualifiers
                1..886
                  /organism="Entamoeba histolytica"

```



```
QY 383 AlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSer 402
Db 831 GCTAAAGCTTCTGCATATAGATTTATCAAAATGATGAATCCACGATATTTGTT 772
QY 403 ArgSerGlnGluLysLysProAspThrCysGluGluGluValPheArgGluVal 422
Db 771 AGATCTACAGCTGGTGAATGTCCTCACTGAGTGTAATGGAATATACATAGAGATGTT 712
QY 423 SerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIle 442
Db 711 CAATTTAGATATCCAAAGACCACTAAACAAATCTTGGTGACCTGATCTTGAAT 652
QY 443 GluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerVal 462
Db 651 AAGAAGAGCAACACAGTTCATAGTAGGAGCATCAGGATGTGGTAAATCAACTACTATT 592
QY 463 GlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAsp 482
Db 591 CAATAGTCCAAAGAAATATGATCCAGTGTGGTGCTCAGTCAATATAGATGGAAGAAT 532
QY 483 AlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluPro 502
Db 531 TTAAGAGATTAAATATCAATGTTAAGAAATCAATAGGATAGTTGGACAGAACCT 472
QY 503 ValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValVal 522
Db 471 ATTTGTTTGGATGTACTATTAGAGAAAATATTATCTTGGAGCTAGAGATGAGAAAACA 412
QY 523 ProLeu--AspGluIleLysGluAlaAlaAsnAlaAsnIleHisSerPheIleGlu 541
Db 411 CCACACTGAAGAGAGATGATGATGCTGCTAAATGGCAATGCAATGAATTTATTTCT 352
QY 542 GlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyValaGlnLeuSerGlyGly 561
Db 351 CATCTTCCAGAGGATATGATCAATGTTAGGAGAAAAGAGCTGCTATATCAGGAGCA 292
QY 562 GlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeu 581
Db 291 CAAAACAAAGAAATGCTATTGTCAGCTGATGATGATGATGATGATGATGATGATGAT 232
QY 582 AspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAsp 601
Db 231 GATGAAGCTACATCAGCATCTGATACAAAGTGAAGAAATGTTACAAACAGCTTGAA 172
QY 602 LysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsn 621
Db 171 AAAGCTAGTCAAGGAAGAACAAATGTTGTAGCACATAGATTAACTGTTAGAAAT 112
QY 622 AlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlu 641
Db 111 GCAAGTAGAAATTTGTGTATTCCTCAAGGAGAAATATTATAA-CAAGGAACACATCAAGAA 53
QY 642 LeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 657
Db 52 TTAAGGATTTAAAGCAACATATATTGATTTATTTTCCAGCAGCAAACT 5

RESULT 13
LOCUS BQ720763 939 bp mRNA linear EST 16-JUL-2002
DEFINITION AGENCOURT_8233170 Lupski_dorsal_root_ganglion Homo sapiens cDNA
clone IMAGE:6184557 5', mRNA sequence.
ACCESSION BQ720763
VERSION BQ720763.1 GI:21859660
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 939)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
```

Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. James R. Lupeki
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation (LLNL)
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LLAM13573 row: 0 column: 22
High quality sequence stop: 646.
Location/Qualifiers
1..939
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6184557"
/clone_lib="Lupski_dorsal_root_ganglion"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev_stage="adult, 36 yr"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6 (Life Technologies); Site_1: Not1; Site_2: SalI; cDNA made by oligo-dT priming. Directionally cloned using the following adaptors: 5'-TCGACCCACGCGTCG-3' and 5'-TCGAGCTTCTAGATCGGAGCGCGCCCT(15)-3'. Size selected > 1 kb for average insert length 1.7 kb. This is a primary library, non-amplified. Library constructed by Life Technologies and donated by J. Lupeki, M.D./Ph.D. (Baylor College of Medicine) and is available through Life Technologies."

BASE COUNT 272 a 202 c 221 g 244 t
ORIGIN

Alignment Scores:
Pred. No.: 1.07e-75 Length: 939
Score: 721.50 Matches: 140
Percent Similarity: 74.44% Conservative: 61
Best Local Similarity: 51.85% Mismatches: 66
Query Match: 21.71% Indels: 3
DB: 14 Gaps: 1

US-09-873-409-1 (1-659) x BQ720763 (1-939)

QY 161 ArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyr 180
Db 3 AAAGCTGGAGAGATCTCCACCAAGCGCTCCGATACATGGTTTCGATCCATGCTCAGA 62
QY 181 GlnAspIleAlaTrpPheAspGlnLysGluAsnSerThrGlyGlyLeuThrThrIleLeu 200
Db 63 CAGGATGTGAGTTGGTTTGTATGATGAGCTAAAAACACCACTGGAGCATTGACTACCAAGGCTC 122
QY 201 AlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGln 220
Db 123 GCCAATGATGCTGCTCAAGTTAAAGGGGCTATAGTTCCAGGCTTGTGTAATACCCAG 182
QY 221 AsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyrGluMetThr 240
Db 183 AATATAGCAATCTGGACAGGAATAATTATATCTTCATCTATGTTGGCACTAACA 242
QY 241 PheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAla 260
Db 243 CTGTACTCTTAGCAATTTGATCCCATCATTCGATGAGGAGTTGTTGAAATGAAATG 302
QY 261 MetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyValIleAlaThr 280
Db 303 TTGCTGGACCAAGCACTGAAAGATAAGAAAGATAGAGGTTCTGGGAAGATCCCTACT 362
QY 281 GluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGln 300
Db 363 GAAGCAATAGAAAACCTCCGAAACCGTTGTTTCTTGATCTCAGGAGCAGAGTTGAAACAT 422
QY 301 MetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIle 320
Db 320


```
Db 423 ATGTATGCTCAGAGTTGTCAGGTACCATACAGAAATCTTTGAGGAAAGCACACATCTTT 482
Qy 321 GlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArg 340
Db 483 GGAATTACATTTCTCCACCAGCAATGATGATATTTTCTCTATGCTGGATGTTCCGG 542
Qy 341 PheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPhe 360
Db 543 TTTGGAGCTACTTGGTGGCACAATAAATCATGAGCTTTGAGGATGTTCTGTAGTATTT 602
Qy 361 ThrAlaIleAlaTyrGlyAlaMetAlaIleGlyThrLeuValLeuAlaProGluTyr 380
Db 603 TCAGCTGTGTCTTGGTGGCCATGGCCGGAAGTCAAGTCAGTTCAATTCCTCCGACTAT 662
Qy 381 Ser-LysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIle 400
Db 663 GCCAAGCCAAATATACAGAGCCACATCATGATCATGATGATGATGATGATGATGAT 722
Qy 400 eAspSerArgSer-GlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGlu---P 419
Db 723 TGACAGCTACAGCACCGAGCCCTAATGCGCAACACATTCGAAGGAAATGTCACCATTT 782
Qy 419 heArgGluValSerPhePheTyrPro 427
Db 783 GGTGAAGTTGTATTTCAACTATCCC 808

RESULT 14
AZ541090 897 bp DNA linear GSS 14-NOV-2000
LOCUS ENTDS67TR Entamoeba histolytica Sheared DNA Entamoeba histolytica
DEFINITION genomic, DNA sequence.
ACCESSION AZ541090
VERSION AZ541090.1 GI:11148493
KEYWORDS GSS.
SOURCE Entamoeba histolytica.
ORGANISM Entamoeba histolytica
Eukaryota; Entamoebidae; Entamoeba.
REFERENCE 1 (bases 1 to 897)
AUTHORS Loftus,B., Van Aken,S. and Fraser,C.
TITLE Determination of clone end sequences from Entamoeba histolytica
JOURNAL HM1:IMSS sheared DNA library
COMMENT Unpublished (2000)
Contact: Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 3543
Email: b.loftus@tigr.org
Clones are derived from the Entamoeba histolytica HM1:IMSS sheared
DNA library
Seq primer: M13-Reverse
Class: shotgun
High quality sequence start: 81
High quality sequence stop: 890.
Location/Qualifiers
1..897
/organism="Entamoeba histolytica"
/strain="HM1:IMSS"
/db_xref="taxon:5759"
/clone_lib="Entamoeba histolytica Sheared DNA"
/note="Vector: pHOSt1, Site 1: Bet 1; Constructed at The
Institute for Genomic Research (TIGR), Rockville, MD.
Genomic DNA isolated from broth cultures of E. histolytica
using a method described by Clark and Diamond (Clark,
C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
method for isolate identification. Exp. Parasitol.
77:450.). The DNA was mechanically sheared to give a
tight size distribution (~2 kb). The v + i method used for
the library construction is described in detail in Smith,
H.O. and Venter, J.C. (Making small insert libraries for
whole genome shotgun sequencing projects. In Genome
Sequencing: A Practical Approach, eds. M. Vaudin and B.
```

```
BASE COUNT 335 a 110 c 177 g 275 t
ORIGIN
Alignment Scores: 1,31e-75 Length: 897
Pred. No.: 720.50 Matches: 143
Score: 74.59% Conservative: 39
Best Local Similarity: 58.61% Mismatches: 61
Query Match: 21.68% Indels: 1
DB: 17 Gaps: 1

US-09-873-409-1 (1-859) x AZ541090 (1-897)
Qy 414 GluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPhe 433
Db 94 AAAGTGAATTTAAAGACATTTGTTTCAGATATCCAAACAGACCAGACAAATCT 153
Qy 434 IleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySer 453
Db 154 GTCTTGAAGGTATTTTCATTTCAAAGTAGAACAGGAAAACTGTTGCAATTAGTAGGAGCA 213
Qy 454 SerCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGln 473
Db 214 TCAGGATGTGGTAAATCAACATCAGTTGATTGAAAGATTTTATGATCCACACAT 273
Qy 474 GlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSer 493
Db 274 GGAGATGATTTATGATGACATAATATCAAGATTTGAATATTCATTTCTTAAGAAGT 333
Qy 494 GlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIle 513
Db 334 CAAATTGGAATGGTAGGACAAAGAACAGTATTTTCTGTAAGTGTATGCGATAATATT 393
Qy 514 AlaTyrGlyAspAsnSerArgVal---ValProLeuAspGluIleLysGluAlaAlaAsn 532
Db 394 AGAAGAGGAGTACCTAAAGAGGTTGAAGTAAGTAATGAACAAATTTATGCTGCTGCTAAA 453
Qy 533 AlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGly 552
Db 454 ATGCAATGCAATGATGATTTATTTTCAGCAATGCCAGAGGATATACACAATGGTAGT 513
Qy 553 LeuLysGlyAlaGlnLeuSerGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeu 572
Db 514 GATAGAGTGCACAAATTTTCAGGAGGACAAACAAAGAAATTTGCTATTCACGTCGCTTG 573
Qy 573 LeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSer 592
Db 574 ATTAGAATCCAAAGTGTATTATTTCTGATGAAGCTACATCAGCACTTGATTCAGAAAGT 633
Qy 593 GluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValVal 612
Db 634 GAAAGATTTGACCAAGATGCACTTGACCAAGCAAAAGGAGGAGGAGGAGGAGGAGGAG 693
Qy 613 ThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLys 632
Db 694 GCATAGATTTCACTATTTCAAAATGCAGATCAAAATATGTTGTTATTCAGAGGAGAGA 753
Qy 633 IleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeu 652
Db 754 ATTCAGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 813
Qy 653 ValAsnAlaGln 656
Db 814 GCTATGCACAA 825

RESULT 15
BH155700
LOCUS BH155700 913 bp DNA linear GSS 24-SEP-2001
DEFINITION ENTRO54TR Entamoeba histolytica Sheared DNA Entamoeba histolytica
ACCESSION BH155700
VERSION BH155700.1 GI:15727822
```


KEYWORDS
Entamoeba histolytica.
SOURCE
Entamoeba histolytica
ORGANISM
Eukaryota; Entamoebidae; Entamoeba.

REFERENCE
1. (bases 1 to 913)
Lofthus B., Wang Z., Van Aken S. and Fraser C.
AUTHORS
Determination of clone end sequences from Entamoeba histolytica
TITLE
HM1:IMSS sheared DNA library (2001)
JOURNAL
Unpublished (2001)

COMMENT
Contact: Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 3543
Email: bjloftus@tigr.org
Clones are derived from the Entamoeba histolytica HM1:IMSS sheared
DNA library
Seq primer: M13-Reverse
Class: shotgun
High quality sequence start: 7
High quality sequence stop: 805.
Location/Qualifiers
1..913
/organism="Entamoeba histolytica"
/strain="HM1:IMSS"
/db_xref="taxon:5759"
/clone_lib="Entamoeba histolytica Sheared DNA"
/note="Vector: pROS1; Site_1: Bst I; Constructed at The
Institute for Genomic Research (TIGR), Rockville, MD.
Genomic DNA isolated from broth cultures of E. histolytica
using a method described by Clark and Diamond (Clark,
C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
method for isolate identification. Exp. Parasitol.
77:450.). The DNA was mechanically sheared to give a
tight size distribution (~2 kb). The v + i method used for
the library construction is described in detail in Smith,
H.O. and Venter, J.C. (Making small insert libraries for
whole genome shotgun sequencing projects. In Genome
Sequencing: A Practical Approach, eds. M. Vaudin and B.
Barell, Oxford University Press, 1999)."

BASE COUNT 359 a 118 c 185 g 251 t

ORIGIN

Alignment Scores:

Pred. No.:	1,64e-74	Length:	913
Score:	711.50	Matches:	146
Percent Similarity:	69.37%	Conservative:	42
Best Local Similarity:	53.87%	Mismatches:	82
Query Match:	21.41%	Indels:	2
DB:	17	Gaps:	1

US-09-873-409-1 (1-659) x BH155700 (1-913)

```

QY   380 TyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsn 399
      ::::: |||||::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
Db   29 TTTTGCTACAGCTAAAGCTTCATCAATAGAATTATCAACAATGGATGAATCCAGAT 88

QY   400 IleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPhe 419
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   89 ATTGATTGTAGATCTACAGCTGGTGAATGTCCAAGTGTGAATGGAATAATTATACATTA 148

QY   420 ArgGluValSerPhePheTyroCysArgProaspValPheIleLeuArgGlyLeuSer 439
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   149 GAAGATGTTCAATTAGATATCCAAACAGACCACCTAACCAAATTTCTGGTGGACTTGAT 208

QY   440 LeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSer 459
      ||||| ::::: ||||| ::::: ||||| ::::: ||||| ::::: ||||| :::::
Db   209 CTTGAAATTAAGAAGGACAAACAGTTGCATTAGTAGAGCATCAGGATGTGGTAAATCA 268

QY   460 ThrSerValGlnLeuGlnArgLeuTyroAspProValGlnGlyGlnValLeuPheAsp 479
      ||||| ::::: ||||| ::::: ||||| ::::: ||||| ::::: ||||| :::::
Db   269 ACTACTATTCAATAGTCCAAAGAAATTTATGATCCAGTTGGTGGGTGAGTCAAAATTAGAT 328

```

Search completed: March 31, 2003, 13:49:30
Job time : 4561.85 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 31, 2003, 02:24:20 ; Search time 96.5072 Seconds
(without alignments)
2094.145 Million cell updates/sec

Title: US-09-873-409-1

Perfect score: 3323

Sequence: 1 MLAEKGAHAELMAKRLGLYS.....QELLRNDRIFYKLVAQSVQ 659

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop .6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09873409/runat 27032003.115422.19259/app_query.fasta_1.7544
-DB=issued Patents NA -QMT=fastcap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09873409 @cgn_1_1_578 @runat 27032003.115422.19259 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:

1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1872	56.3	2726	1 US-08-461-823-1	Sequence 1, Appli
2	1872	56.3	4264	2 US-08-784-649A-1	Sequence 1, Appli
3	1872	56.3	4264	2 US-08-784-649A-5	Sequence 5, Appli
4	1872	56.3	4646	1 US-08-181-471-2	Sequence 2, Appli
5	1872	56.3	4669	6 5206352-3	Patent No. 5206352
6	1872	56.3	6505	2 US-08-793-610-5	Sequence 5, Appli
7	1872	56.3	9318	2 US-08-793-610-6	Sequence 6, Appli
8	1855	55.8	4669	2 US-08-752-447-1	Sequence 1, Appli
9	1855	55.8	4669	4 US-09-316-167-1	Sequence 1, Appli
10	1849.5	55.7	4233	3 US-09-120-513-1	Sequence 1, Appli
11	1849.5	55.7	4233	4 US-09-450-105-1	Sequence 1, Appli
12	1845	55.5	4669	2 US-08-583-276-18	Sequence 18, Appl

13	1254	37.7	4002	2	US-08-996-545-1	Sequence 1, Appli
14	1254	37.7	4002	2	US-08-996-545-3	Sequence 3, Appli
15	1254	37.7	4002	4	US-09-328-320-1	Sequence 1, Appli
16	1254	37.7	4002	4	US-09-328-320-3	Sequence 3, Appli
17	1227.5	36.9	4047	2	US-08-612-734B-1	Sequence 1, Appli
18	1184.5	35.6	4800	2	US-08-612-734B-3	Sequence 3, Appli
19	1151	34.6	3924	1	US-08-395-246C-1	Sequence 1, Appli
20	1049.5	31.6	4224	1	US-08-612-521-1	Sequence 1, Appli
21	975	29.3	3909	1	US-08-232-537-1	Sequence 1, Appli
22	851	25.6	6143	1	US-08-612-521-3	Sequence 3, Appli
23	842	25.3	3924	2	US-08-996-644-3	Sequence 3, Appli
24	842	25.3	3924	3	US-09-352-552-3	Sequence 3, Appli
25	842	25.3	3927	2	US-08-996-644-1	Sequence 1, Appli
26	842	25.3	3927	3	US-09-352-552-1	Sequence 1, Appli
27	840	25.3	3792	4	US-09-351-224E-10	Sequence 10, Appli
28	783	23.6	3999	4	US-09-351-224E-9	Sequence 9, Appli
29	677	20.4	2376	1	US-08-394-880B-1	Sequence 1, Appli
30	644	19.4	2244	4	US-09-061-764A-18	Sequence 18, Appli
31	623.5	18.8	2061	4	US-09-061-764A-17	Sequence 17, Appli
32	608.5	18.3	5120	3	US-08-772-270A-6	Sequence 6, Appli
33	608.5	18.3	8370	2	US-08-488-706-1	Sequence 1, Appli
34	601	18.1	1749	4	US-09-134-001C-1893	Sequence 1893, Ap
35	600.5	18.1	1743	4	US-09-134-001C-2774	Sequence 2774, Ap
36	599	18.0	13188	4	US-08-961-527-70	Sequence 70, Appli
37	594.5	17.9	7721	3	US-08-772-270A-14	Sequence 14, Appli
38	555.5	16.7	1959	4	US-09-061-764A-4	Sequence 4, Appli
39	555.5	16.7	7760	4	US-08-961-527-63	Sequence 63, Appli
C 40	552.5	16.6	28804	2	US-08-592-874-1	Sequence 1, Appli
C 41	552.5	16.6	28804	3	US-09-096-942-2	Sequence 2, Appli
C 42	552.5	16.6	28804	3	US-09-096-867-2	Sequence 2, Appli
43	548	16.5	6492	4	US-08-961-527-188	Sequence 188, Appli
44	544.5	16.4	4411529	4	US-09-103-840A-1	Sequence 1, Appli
45	543.5	16.4	7186	4	US-08-961-527-39	Sequence 39, Appli

ALIGNMENTS

RESULT 1

US-08-461-823-1
; Sequence 1, Application US/08461823
; Patent No. 5593840
; GENERAL INFORMATION:
; APPLICANT: Bhatnagar, Satish K.
; APPLICANT: George Jr., Albert L.
; APPLICANT: Nazarenko, Irina
; TITLE OF INVENTION: AMPLIFICATION OF NUCLEIC ACID SEQUENCES
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OncorPharm, Inc.
; STREET: 200 Perry Parkway
; CITY: Gaithersburg
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20877
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461.823
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/168,621
; FILING DATE: 16-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/010,433
; FILING DATE: 27-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Karta, Glenn E.
; REGISTRATION NUMBER: 30,649
; REFERENCE/DOCKET NUMBER: PA-0012 CIP 2

TELECOMMUNICATION INFORMATION:

TELEPHONE: 301 527-2058
 TELEFAX: 301 208-6997
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2726 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 US-08-461-823-1

Alignment Scores:

Pred. No.: 2,8e-219 Length: 2726
 Score: 1872.00 Matches: 356
 Percent Similarity: 74.22% Conservative: 142
 Best Local Similarity: 53.06% Mismatches: 159
 Query Match: 56.33% Indels: 14
 DB: 1 Gaps: 4

US-09-873-409-1 (1-659) x US-08-461-823-1 (1-2726)

Qy 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
 Db 317 GTCATTGTGAGAAAGGAATCATGATGAACATCATGAAAGAGAAAGGCATTTACTTCAA 376
 Qy 21 LeuValMetSerGln-----AspIleLysAlaAspGluMet 34
 Db 377 CTTGTCAACATCGACAGCAGGAAATGAAGTTGAATTAGAAATGCAGCTGATGAATCC 436
 Qy 35 GluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--- 53
 Db 437 AAAAGTGAATGATGCGCTTGGAAATGCTTCAATGATTCAGATCCAGTCTAATAAGA 496
 Qy 54 -----LysSerIleLys---SerAspPheIleAspLysAlaGluLys 67
 Db 497 AAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAAGCCCAAGCAAGCTTAGTACC 556
 Qy 68 ThrGlnSerLysGluLysSerLeuProGluValSerLeuLysIleLeuLysLeuAsn 87
 Db 557 AAAGAGGCTCTGGATGAATGATATACCTCCAGTTCTCTTTGGAGGATTAATGAAGCTAAT 616
 Qy 88 LysProGluTyrProPheValValLysGlyThrLeuAlaSerValLeuAsnGlyThrVal 107
 Db 617 TTAACGTAATGCGCTTATTTGTGTGTGTATTTGTGCGCATATATAATGGAGCGCTG 676
 Qy 108 HisProValPheSerIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAsp 126
 Db 677 CAACCGACATTTGCAATAATATTTTCAAGATATATAGGGGTTTTTACAAAGAATTGATGAT 736
 Qy 127 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 146
 Db 737 CCGTAAACAAACACGACAGATATGTAATCTTGTCTACTATTGTTCTAGCCCTTGGAAAT 796
 Qy 147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 166
 Db 797 ATTTCTTTTATACATTTTCTTCCGTTTCCAGGTTTCCATTTTGGCAAGCTGGAGATCTTC 856
 Qy 167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPhe 186
 Db 857 ACCAAGCGCTCCGATACATGTTTCCGATCCATGCTCAGACAGGATGTGATGTTGTTT 916
 Qy 187 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206
 Db 917 GATGACCTTAAACACCACTCGAGCATTTGACTTACCAGGCTCGCCCAATGATCTGCTCAA 976
 Qy 207 IleGlnGlyValThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226
 Db 977 GTTAAAGGGGCTATAGTTTCCAGGCTTGTGTAAATTTACCAGCAATATAGCAATCTTGG 1036
 Qy 227 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle 246

Db 1037 ACAGGAATAATTATATCTTCTATGTTGGCAACTAAACACTGTTACTCTTACCAATT 1096
 Qy 247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266
 Db 1097 GTACCATCATTTGCATAGCAGGAGTTGTTGAAATGAAATGTTGCTCGGACAGCACTG 1156
 Qy 267 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286
 Db 1157 AAAGATAAGAAAGAACTAGAAAGTCTCGGAAGATCGCTACTGAAGCAATAGAAAACTTC 1216
 Qy 287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306
 Db 1217 CGAACCGTTGTTCTTCTGACTCAGCAGCAGAAAGTTTGAACATATATGCTCAGAGTTTG 1276
 Qy 307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326
 Db 1277 CAGGTACCATCAGAAATCTCTTTGAGAAAGCACACATCTTTGGAAATTACATTTTCCCTTC 1336
 Qy 327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyValaTyrLeuIle 346
 Db 1337 ACCCAGGCAATGATGATTTTCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTG 1396
 Qy 347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366
 Db 1397 GCACATAAACTCATGAGCTTTGAGGATGTTCTGTTAGTATTTTTCAGCTGTTGCTTTGGT 1456
 Qy 367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386
 Db 1457 GCATGGCGCGTGGGCAAGTCAGTTTCATTTGCTCTGATATGCCCAGCAAAATATCA 1516
 Qy 387 AlaAlaHisLeuPheAlaLeuLeuLysLysProAsnIleAspSerArgSerGlnGlu 406
 Db 1517 GCAGCCCATCATCATCATCATGATCATTTGAAACCCCTTTGATTGACAGCTACACGCGAA 1576
 Qy 407 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPheTyr 426
 Db 1577 GGCCTAATGCGGAAACACATTTGGAAGAAATGTCAATTTGGTGAAGTTGTATTCAACTAT 1636
 Qy 427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446
 Db 1637 CCCACCCAGCGGACATCCAGTGTCTCAGGACCTGAGCCTGGAGTGAAGAGGCGCAG 1696
 Qy 447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuGln 466
 Db 1697 AGCTGGCTCTGTTGGGCGAGCTGGCTGTGGGAAGACACAGTGTCCAGCTCTCTGGAG 1756
 Qy 467 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 486
 Db 1757 CGGTTCTACGACCCCTTGGCAGGAAAGTGTCTGTATGGCAAGAAATAAAGCGACTG 1816
 Qy 487 AsnValGlnTyrLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506
 Db 1817 AATGTTCTAGTGGCTCCGAGCACCTGGGCTGTTGCCAGGACCCATCTCTGTTTAC 1876
 Qy 507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 526
 Db 1877 TGCAGATTGCTGAGAAATTTGCCCTATGGAGACAACACGCGGCTGTGTACAGGAGAG 1936
 Qy 527 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546
 Db 1937 ATCTGTGAGGCGAGCAAGAGGCGCAACATCATGCTTTCATCGAGTCACTGCTTAATAA 1996
 Qy 547 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnIleGlnArgLeu 566
 Db 1997 TATAGCATTAAGTAGGAGACAAGAGAACTCAGCTCTCTGTTGGCCAGAAACAGCAATT 2056
 Qy 567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 586
 Db 2057 GCATAGCTCGTGGCTTGTAGACAGCTCATATTTTGTCTTTTGTGATGAGCCAGCTCA 2116
 Qy 587 AlaLeuAspAsnSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606

Db 2117 GCTCTGGATACAGAAAGTGAAGGTTGTCTCAAGAGCCCTGGACAAAGCCAGCAGAGGCC 2176
Qy 607 ArgThrCysLeuValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626
Db 2177 CGCACCTGCAATGTGTGCTCACCGCTGTCCACCATCCAGATGCAGACTTAATAGTG 2236
Qy 627 ValLeuHisAsnGlyLysIleGlyGluGlnGlyThrHisGlnGlnLeuLeuLeuArgAsnArg 646
Db 2237 GTGTTTCAGATGCGCAGAGTCAAGGAGCATGCGCAGCATCAGCAGCTGCTGGCACAGAAA 2296
Qy 647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657
Db 2297 GGCATCTATTTTCAATGGTCAAGTGTCCAGGCT 2329
RESULT 2
US-08-784-649A-1
; Sequence 1, Application US/08784649A
; Patent No. 5830697
; GENERAL INFORMATION:
; APPLICANT: Sikic, Branimir I
; APPLICANT: Chen, Gang
; TITLE OF INVENTION: P-GLYCOPROTEIN MUTANT RESISTANT TO
; TITLE OF INVENTION: CYCLOSPORIN MODULATION
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 2200 Sand Hill Road
; CITY: Menlo Park
; STATE: CA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/784,649A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: Reg.No. 5830697 36,677
; REFERENCE/DOCKET NUMBER: 06037/007001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-322-5070
; TELEFAX: 415-854-0875
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4264 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-784-649A-1
Alignment Scores:
Pred. No.: 6,17e-219 Length: 4264
Score: 1872.00 Matches: 356
Percent Similarity: 74.22% Conservative: 142
Best Local Similarity: 53.06% Mismatches: 159
Query Match: 56.33% Indels: 14
DB: 2 Gaps: 4
US-09-873-409-1 (1-659) x US-08-784-649A-1 (1-4264)
Qy 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
Db 1951 GTCATTGTGGAGAAAGAAATCATGTAACATCATGAAAGAGAAAGGCAATTTACTTCAA 2010
Qy 21 LeuValMetSerGln-----AspIleLysLysAlaAspGluGlnMet 34
Db 2011 CTTGTCAATGTCAGACAGCAGGAAATGAAGTTGAATTAGAAATGTCAGCTCATGAATCC 2070

Qy 35 GluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--- 53
Db 2071 AAAAGTGAATGATGCTTGGAAATGCTCTCAATGATTAAGATCAAGATCCAGTCTAATAAGA 2130
Qy 54 -----LysSerIleLys---SerAspPheIleAspLysAlaGluGluSer 67
Db 2131 AAAAGATCAACTGATGAGTGTCCTGGGATCAACGCCNAGACAGAAAGCTTAGTACC 2190
Qy 68 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 87
Db 2191 AAGAGGCTCTGGATGAAGATATACCTCCAGTTTCTTTTGGAGGATATAGAGCTTAAT 2250
Qy 88 LysProGluTyrPheValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107
Db 2251 TTAACCTGAATGGCTTATTTTGTGGTGTATTTTGTGTCATTATAAATGGAGGCTG 2310
Qy 108 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAsp 126
Db 2311 CAACGACATTTGCAATAATATTTCAAGATATATAGGGTTTTACAGAAATGTATGAT 2370
Qy 127 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 146
Db 2371 CCTGAACAAACACGACAGATAGTAAGTCTTCTTCTTCTAGCCCTTGGAAAT 2430
Qy 147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 166
Db 2431 ATTTCTTTTATATACATTTTCTTCCAGGTTTTCATTTTGGCAAGCTGGAGAGATCCTC 2490
Qy 167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPhe 186
Db 2491 ACCAAGCGCTCCGATACATGTTTCCGATCCATGCTCAGACAGAGATGTGATGGTTT 2550
Qy 187 AspGluLysGluAsnSerThrGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206
Db 2551 GATCACCTTAAACACACCATGGAGCATTCATACAGGCTCGCAATGATGCTCTCAA 2610
Qy 207 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226
Db 2611 GTTAAAGGGCTATAGGTTCCAGGCTTGTCTGTAATACCCAGAATATAGCAAAATCTTGG 2670
Qy 227 LeuSerValIleIleSerPheIleTyrGlyThrGluMetThrPheLeuIleLeuSerIle 246
Db 2671 ACAGAAATATATATCTTCTTCTATGTTGGCAATACACATGTTACTCTTAGCAAT 2730
Qy 247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266
Db 2731 GTACCATCATGCAATAGCAGGAGTTGTTGAATGAAATGTTGCTGGCAAGCACTG 2790
Qy 267 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286
Db 2791 AAGATAAGAAAGAACTAGAAAGTGTGGGAAGATCGCTACTGAAGCAATAGAAAACCTC 2850
Qy 287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306
Db 2851 CGAACCGTTGTTCTTCTTCTGACTCAGAGCAGAAAGTTTGAACATATGATGCTCAGAGTTG 2910
Qy 307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysThrAlaPhe 326
Db 2911 CAGGTACCATACAGAACTCTTTGAGGAAAGCACACATCTTTGGAAATTTACATTTCTTC 2970
Qy 327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrIleuIle 346
Db 2971 ACCCAGGCAATGATGATTTTCTATGCTGATGTTTCCGGTTTCCGGTCTTGGAGCTACTTGG 3030
Qy 347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366
Db 3031 GCACATAAATCATGAGCTTTGAGGATGTTCTGTAGTATTTTTCAGCTTTGTCTTCTTGT 3090
Qy 367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386
Db 3091 CCATGGCCGTTGGGCAAGTCAGTTCATTTCTGCTATGCTATGCCAAAGCCCAANAATCA 3150


```
QY 387 AlaAlaHisLeuPheAlaLeuLeuGluLysLeuProAsnIleAspSerArgSerGlnGlu 406
Db 3151 GCAGCCACATCATCATGATCAATGAAACCCCTTTGATTGACAGCTACACAGGAA 3210
QY 407 GlyLysLeuProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPheThr 426
Db 3211 GGCTTAATGCGGACACATTTGGAGGAATGTCACATTTGGTGAAGTTGATTCAACTAT 3270
QY 427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446
Db 3271 CCACCGGACCGACATCCAGTCTTCAGGACTGAGCCCTGAGGTGAAGAAGGCCAG 3330
QY 447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 466
Db 3331 ACCTGCTCTGCTGGCAGCAGTGGCTGGGAGAGACACAGTGGTCCAGCTCTCTGGAG 3390
QY 467 ArgLeuThrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 486
Db 3391 CGTTTCTACAGCCCTTTGGCAGGGAAGTCTGCTTGTATGCCAAGAAATAAAGCGACTG 3450
QY 487 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506
Db 3451 AATGTCAGTGGCTCCAGACCATCCACCTGGGATCGTGTCCAGGAGCCCATCTCTTTGAC 3510
QY 507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 526
Db 3511 TGCAGCATTTGCTGAGAACATTTGCTTATGGAGACAACAGCCGGTGTGTCCACAGGAAGAG 3570
QY 527 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546
Db 3571 ATTGTCAGGCGACCAAGGAGGCGCAACATACATGCTTCATCGAGTCACTGCGCTAATAA 3630
QY 547 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlnLysGlnArgLeu 566
Db 3631 TATAGCACTAAGTAGAGCAAGAAAGTGTTCCTGCTGCGCCAGAAACACCGCAT 3690
QY 567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 586
Db 3691 GCCATAGTCTGCTGCTTGTAGACAGCTCATATTTTGTGTTGGATGAAGCCACGTCA 3750
QY 587 AlaLeuAspAsnAspSerGluLysValGlnHisAlaLeuAspLysAlaArgThrGly 606
Db 3751 GCTCTGGATACAGAAAGTGAAGAGTGTTCCTCAAGAGCCCTGGACAAAGCCAGAGAGGC 3810
QY 607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626
Db 3811 CGCACTGCAATGTTGATGCTCAACCCCTGTCACCATCCAGATCCAGATCACTTAATAGTG 3870
QY 627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 646
Db 3871 GTGTTTTCAGAAATGGCAGAGTCAAGGAGCATGGCAGCATCAGCAGCTGCTGGCAGAGAA 3930
QY 647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657
Db 3931 GGCATCTATTTTCAATGGTCAAGTGTCCAGGCT 3963
```

RESULT 3

US-08-784-649A-5

; Sequence 5, Application US/08784649A

; Patent No. 5830697

; GENERAL INFORMATION:

; APPLICANT: Sikic, Branimir I

; APPLICANT: Chen, Gang

; TITLE OF INVENTION: P-GLYCOPROTEIN MUTANT RESISTANT TO

; NUMBER OF INVENTIONS: CYCLOSPORIN MODULATION

; NUMBER OF SEQUENCES: 5

; CORRESPONDENCE ADDRESS:

; ADDRESSER: Fish & Richardson

; STREET: 2200 Sand Hill Road

; CITY: Menlo Park

; STATE: CA

; COUNTRY: USA

; ZIP: 94025

```
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/784,649A
FILING DATE:
```

```
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sherwood, Pamela J
REGISTRATION NUMBER: Reg.No. 5830697 36,677
REFERENCE/DOCKET NUMBER: 06037/007001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-322-5070
TELEFAX: 415-854-0875
```

```
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 4264 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-784-649A-5
```

```
Alignment Scores:
Pred. No.: 6,17e-219 Length: 4264
Score: 1872.00 Matches: 356
Percent Similarity: 74.22% Conservative: 142
Best Local Similarity: 53.06% Mismatches: 159
Query Match: 56.33% Indels: 14
DB: 2 Gaps: 4
```

US-09-873-409-1 (1-659) x US-08-784-649A-5 (1-4264)

```
QY 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrSer 20
Db 1951 GTCAATTGTGGAGAAAGGAATCATGATGAACCTCATGAAAGAGAGAGGCACTTTACTTCAA 2010
QY 21 LeuValMetSerGln-----AspIleLysIleAlaAspGluGlnMet 34
Db 2011 CTTGTCAACAATCAGACAGACAGGAATGAAGTTGAATTAAGAAATGCAGCTGATGAATCC 2070
QY 35 GluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--- 53
Db 2071 AAAAGTGAATTTGATGGCTTGGAAATGCTTCAATATGATTCAGATCCAGCTCTAATAGA 2130
QY 54 -----LysSerIleLys---SerAspPheIleAspLysAlaGluGluSer 67
Db 2131 AAAAGATCAACTCGTAGGAGTGTCCGTGGATCAAGCCCAAGACAGAAAGCTTAGTACC 2190
QY 68 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLysIleLeuLysLeuAsn 87
Db 2191 AAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCTTTTGGAGGATTAATGAAGCTAAAT 2250
QY 88 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107
Db 2251 TTAACGTGAATGCTTATTTTGTGTGTATTTTGTGCTATTAATATGAAGCGCTG 2310
QY 108 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAsp 126
Db 2311 CAACAGCATTTTGAATAATATTTTCAAGATATATAGGGGTTTTTACAAGAAATGATGAT 2370
QY 127 LysThrThrLeuLysHisAspAlaGluIleLysSerMetIlePheValIleLeuGlyVal 146
Db 2371 CCTGAACAAAACAGACAGAAATAGTAACCTGTTTCTACTATTGTTCTAGCCCTTGAAT 2430
QY 147 IleCysPheValSerPheMetGlnGlyLeuPheThrGlyArgAlaGlyGluIleLeu 166
Db 2431 ATTTCTTTTATACATTTTCTCTCAGGGTTTCACATTTGGCAAGAGCTGAGAGATCCTC 2490
QY 167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 186
```


Dbb 2491 ACCAAGCGGCTCCGATACATGCTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTT 2550
Qy 187 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206
Dbb 2551 GATGACCTTAAACACACCACTGGAGCAATGACTACCAAGCTCGCCAATGATGCTCTCAA 2610
Qy 207 IleGlnGlyAlaThrGlySerArgIleGlyValLeuLeuThrGlnAsnAlaThrAsnMetGly 226
Dbb 2611 GTTAAAGGGCTATAGGTTCCAGGCTTGCTGTAATATACCAAGATATAGCAAAATCTGGG 2670
Qy 227 LeuSerValIleIleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIle 246
Dbb 2671 ACAGAAATATATATCTTCACTATGCTGGCAACTAACACTGTTACTCTTACCAATT 2730
Qy 247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266
Dbb 2731 GTACCCATCATTCATAGCAGGATGTTGAAATGAAATGTTGCTGGACAAGCACTG 2790
Qy 267 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286
Dbb 2791 AAGATAAGAAAGACTAGAGGCTGGGAAGATCGCTACTGAAGCAATAGAAAACCTTC 2850
Qy 287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluLumMetLeu 306
Dbb 2851 CGAACCGTGTGTTCTTGAATCAGGAGCAGAAAGTTGAACATATATGATGCTCAGAGTTG 2910
Qy 307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326
Dbb 2911 CAGGTACCATACAGAAACTCTTTGAGGAAGCACACATCTTTGGAAATACATTTCCCTTC 2970
Qy 327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 346
Dbb 2971 ACCCAGGCAATGATGATATTTTCTATGCTGATGTTCCGGTTGGAGCCTACTTGGTG 3030
Qy 347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366
Dbb 3031 GCACATAAATCACTCAGCTTTGAGGATGTTCTGTAGTATTTTCAGCTGTTCTCTTGTG 3090
Qy 367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386
Dbb 3091 GCATGGCGGTGGGCAAGTCAGTTCATTTGCTCTGACTATGCCAAGCCAAATATCA 3150
Qy 387 AlaAlaHisLeuPheAlaLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 406
Dbb 3151 CGAGCCACATCATGATGATCATTTGAAAAACCCCTTTGATTCAGACGTACACAGCAGGAA 3210
Qy 407 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 426
Dbb 3211 GGCCTAATGCCCAACACATTTGGAAGGAATGTGCATTTTGGTGAAGTTGTATTCAACTAT 3270
Qy 427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446
Dbb 3271 CCCACCGACCGGACATCCAGTGCTTCAGGAGCTGAGCTCGAGTGAAGAGGCCAG 3330
Qy 447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuGln 466
Dbb 3331 ACCTGGCTCTGCTGGCAGCAGTGGCTGTGGGAAGACACAGTGTCTCCAGCTCTCTGGAG 3390
Qy 467 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 486
Dbb 3391 CGGTTCACGACCCCTTGGCAGGGAAGTGTCTGTATGGCAAGAAATAAAGGCACTG 3450
Qy 487 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506
Dbb 3451 AATGTTCACTGGCTCGAGACACCTGGGCACTGCTCCAGAGCCCACTCTGTTGAC 3510
Qy 507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 526
Dbb 3511 TGCAGATTGCTGAGAACATTCCTATGAGACACACAGCCGGGTGTGTACAGGAAGAG 3570
Qy 527 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546
Dbb 3571 ATTGTGAGGCGCAAGAGGCGCAACATACATACCTTCTTCACTGAGTCACTGCTTAATAA 3630

Qy 547 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 566
Dbb 3631 TATAGCACTAAGTAGGAGACAAAGGAACCTCAGCTCTCTGTTGGCCAGAAACACGCAAT 3690
Qy 567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 586
Dbb 3691 GCCATAGCTGTGCGCTTGTAGACAGCTCATATTTTGTCTTTGGATGAAGCCACGTC 3750
Qy 587 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606
Dbb 3751 GCTCTGGATACAGAAAGTGAAGGTTGTCCAAGAGCCCTGGACAAAGCCAGAGAAGGC 3810
Qy 607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626
Dbb 3811 CGCACTCATTTGATGCTCAGCGCTGTCCACCATCCAGAAATGCAGACTTAATAGTG 3870
Qy 627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuArgAsnArg 646
Dbb 3871 GTGTTTCAGATGCGCAGTCAAGGATGCGCAGCATGCGCAGCTGCTGCGCACAGAA 3930
Qy 647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657
Dbb 3931 GCATCTATTTTCAATGCTCAGTCTCAGGCT 3963
RESULT 4
US-08-181-471-2
; Sequence 2, Application US/08181471
; Patent No. 5641508
; GENERAL INFORMATION:
; APPLICANT: Li, Lingna
; APPLICANT: Lishko, Valeryi K.
; TITLE OF INVENTION: METHOD FOR DELIVERING BENEFICIAL
; TITLE OF INVENTION: COMPOSITIONS TO HAIR FOLLICLES
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Thomas Fitting
; STREET: 12526 High Bluff Drive, Suite 300
; CITY: San Diego
; STATE: CA
; COUNTRY: USA
; ZIP: 92130
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/181,471
; FILING DATE: 13-JAN-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/041,553
; FILING DATE: 02-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: ANT0029P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-792-3680
; TELEFAX: 619-792-8477
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4646 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 425..4267

US-08-181-471-2

Alignment Scores:

Pred. No.: 7,18e-219 Length: 4646
Score: 1872.00 Matches: 356
Percent Similarity: 74.22% Conservative: 142
Best Local Similarity: 53.06% Mismatches: 159
Query Match: 56.33% Indels: 14
DB: 1 Gaps: 4

US-09-873-409-1 (1-659) x US-08-181-471-2 (1-4646)

QY 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrSer 20
DB 2237 GTCAATTGGAGAAAGGAATCATGATGAAGTTCATGAAAGAGAAAGCATTTACTTCAAA 2296
QY 21 LeuValMetSerGln-----LysSerIleLys-----AspIleLysLysAlaAspGluGlnMet 34
DB 2297 CTTGTCACAATGCAGACAGCAGGAATGAAGTTGAATTTAGAAATGCAGCTGAATGAATCC 2356
QY 35 GluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--- 53
DB 2357 AAAAGTCGAATTCATGCTTGGAAATGCTCTCAATGATTCAGATCCAGCTCAATAAGA 2416
QY 54 -----LysSerIleLys---SerAspPheIleAspLysAlaGluSer 67
DB 2417 AAAAGATCAACTCGTAGGAGTGTCCGTGGATCAACAAGCCCAAGACAGAAAGCTTAGTAA 2476
QY 68 ThrGlnSerLysGluLeuSerLeuProGluValSerLeuLeuLysIleLeuLysLeu 87
DB 2477 AAGAGGCTCTGGATGAAAGTATACCTCCAGTTCTCTTTGGAGGATTATGAAGCTAAAT 2536
QY 88 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107
DB 2537 TTAACCTGAATGGCTTATTTGTGTGGTGATTTGTGGCATTATTAATGAGGCGCTG 2596
QY 108 HisProValPheSerIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAsp 126
DB 2597 CAACCAGCATTTTGCATATATATTTTCAAGATTTATAGGGGTTTTTACAAGAAATGATGAT 2656
QY 127 LysThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 146
DB 2657 CCTGAAACAAACGACAGAAATAGTATCTGTCTTCTACTATTTGTCTAGCCCTTGGAAAT 2716
QY 147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 166
DB 2717 ATTTCTTTTATTACATTTTCTTCAGGTTTCCACATTTGGCAAGCTGGAGAGATCCTC 2776
QY 167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 186
DB 2777 ACCAAGGGCTCCGATACATGTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTT 2836
QY 187 AspGluLysGluAsnSerThrGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206
DB 2837 GATGACCCCTAAACACCACTGGAGCATTGACTACCAGGCTCGCCCAATGATGCTGCTCAA 2896
QY 207 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226
DB 2897 GTTAAAGGGCTATAGTTCAGGCTTGTGTAATTTACCAGATATAGCAATCTTGGG 2956
QY 227 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle 246
DB 2957 ACAGGAATAATTATATCTTCATCTATGTTGGCAACTAACACTGTACTCTTAGCAATT 3016
QY 247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266
DB 3017 GTACCCATCATTCATAGCAGAGTGTGTGAATTTGAATTTGTCTGGACAAAGCAGCTG 3076
QY 267 LysAspLysGlnGluLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286
DB 3077 AAAGATAAGAAAGAACTAGAGGTGCTGGAGATCGCTACTAGCAACATAGAAACTTC 3136
QY 287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306

DB 3137 CGAACCGTTGTTTCTTGTGACTCAGGACGAGAAGTTTGAACATATGTATGCTCAGAGTTTG 3196
QY 307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326
DB 3197 CAGGTACCATACAGAAATCTTTTGGAGAAAGCACATCTTTGGAAATACATTTTCTTC 3256
QY 327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 346
DB 3257 ACCCAGGCAATGATGATTTTCTTATGCTGATGTTCCGGTTTGGAGCTTACTTGGTG 3316
QY 347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366
DB 3317 GCACATAAATCATGAGCTTTCAGATGTTCTGTAGTATTTTTCAGCTGTGTCTTGTGT 3376
QY 367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386
DB 3377 GCATGGCGCTGGGCAAGTCACTTTCCTCTGCTACTATGCCAAGCCCAATATATCA 3436
QY 387 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 406
DB 3437 CGAGCCCATCATCATGATCATTTGAAAAAACCCCTTTGATTGACAGCTTACAGCAGGAA 3496
QY 407 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 426
DB 3497 GGCCTAATGCCCAACACATTTGAAGGAATGTACATTTTGGTGAAGTTGTATTCAACTAT 3556
QY 427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446
DB 3557 CCCACCGACCCGAGATCCAGTGTCTTGGGAGCTGAGCTGGAGTGAAGAGGCGCCAG 3616
QY 447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuGln 466
DB 3617 ACGTGGCTCTGTGGCAGCAGTGTCTGTGGAAAGAGCACAGTGTCTCCAGCTCTCTGGAG 3676
QY 467 ArgLeuTyrAspProValGlnGlnValLeuPheAspGlyValAspAlaLysGluLeu 486
DB 3677 CGGTTCTACGACCCCTTGGCAGGAAAGTGTCTGTGTGTCGCAAAATAAAGGAGCTG 3736
QY 487 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506
DB 3737 AATGTCAGTGGCTCGAGCACCTGGGCATCGTGTCCAGAGCCCATCTCTGTTTAC 3796
QY 507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValProLeuAspGlu 526
DB 3797 TGCAGATTGTCTGAGAACATTCCTTATGGAGACAAACAGCCGGTGTGTGTACAGAGAG 3856
QY 527 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546
DB 3857 ATCGTGAGGGCAGCAAGAGGCGCAACATACATGCTTTCATCGAGTCACTGCTTAATAA 3916
QY 547 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 566
DB 3917 TATAGCATTAAGTAGGAGACAAAGAACTCAGCTCTCTGTGTGGCCAGAAACAGCAT 3976
QY 567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuAspGluAlaThrSer 586
DB 3977 GCCATAGCTCGTGGCTTGTAGACAGCTCATATTTTGTCTTGTGTAAGCCACGTCA 4036
QY 587 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606
DB 4037 GCTCTGGATACAGAAAGTGAAGGTTGTCCAAGAAAGCCCTGGACAAAGCCAGAGAGC 4096
QY 607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626
DB 4097 CGCACTGATGATGATGCTCCCGCTGTCCACATTCAGATTCAGACTTAATAGTG 4156
QY 627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuArgAsnArg 646
DB 4157 GTGTTTCAGAAATGGCAGATCAAGGAGCATGGCACGCATCAGCAGCTGTGTGCACAGAAA 4216
QY 647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657

Db 4217 GGCATCTATTTTCAATGGTCAGTCCAGGCT 4249

```
RESULT 5
5206352-3
; Patent No. 5206352
; APPLICANT: Roninson, Igor B.; Pastan Ira H.; Gottesman,
; Michael M.
; TITLE OF INVENTION: COMPOSITIONS FOR CLONES CONTAINING DNA
; SEQUENCES ASSOCIATED WITH MULTIDRUG RESISTANCE IN HUMAN CELLS
; NUMBER OF SEQUENCES: 4
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/622,836
; FILING DATE: 24-SEP-1990
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 892,575
; FILING DATE: 01-AUG-1986
; APPLICATION NUMBER: 845,610
; FILING DATE: 28-MAR-1986
; SEQ ID NO:3:
; LENGTH: 4669
5206352-3
```

```
Alignment Scores:
Pred. No.: 7,24e-219 Length: 4669
Score: 1872.00 Matches: 356
Percent Similarity: 74.22% Conservative: 142
Best Local Similarity: 53.06% Mismatches: 159
Query Match: 56.33% Indels: 14
DB: 6 Gaps: 4
```

US-09-873-409-1 (1-659) x 5206352-3 (1-4669)

```
Qy 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrSer 20
Db 2237 GTCATTGGAGAAAGGAATCATGATGAACATCATGAAGCAAGCAAGGCAATTTACTTCAA 2296
Qy 21 LeuValMetSerGln-----AspIleLysLysAlaAspGluGlnMet 34
Db 2297 CTTGTCACAATCGACAGCAGCAGGAATGAAGTTGAATTAGAAAATGCAGCTCATGAATCC 2356
Qy 35 GluSerMetThrTyr-SerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--- 53
Db 2357 AAAAGTGAATTTGATGCCTTGGAAATGTCTTCAATGATTCAGATCCAGTCTAATAAGA 2416
Qy 54 -----LysSerIleLys---SerAspPheIleAspLysAlaGluGluSer 67
Db 2417 AAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAGCCACAGACAGAAAGCTTAGTACC 2476
Qy 68 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 87
Db 2477 AAAGAGGCTCTGGATGAAGATATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAAT 2536
Qy 88 LysProGluTyrProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107
Db 2537 TTAACCTGAATGSCCTTATTTGTGTGTGTGTTTGTGTCATTTTGGCCATTATAAATGGAGCCTG 2596
Qy 108 HisProValPheSerIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAsp 126
Db 2597 CAACAGCAATTTGCAATATATTTTCAAGATATATAGGGGTTTTTACAGAAATGATGAT 2656
Qy 127 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 146
Db 2657 CCTGAAACAAAACGACAGAAATAGTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAAAT 2716
Qy 147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 166
Db 2717 ATTTCCTTTTATTACATTTTCTTCCAGGGTTTCACATTTTGGCAAGCTCGAGAGATCTCTC 2776
Qy 167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPhe 186
Db 2777 ACCAAGCGGCTCGATACATGTTTTCGATCCATGCTCAGACAGGATGTGAGTTGGTTT 2836
Qy 187 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206
```

```
Db 2837 GATGACCTAAAAACACACACTGGAGCATTTGACTACAGGCTCGCAATGATGCTGCTCAA 2896
Qy 207 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226
Db 2897 GTTAAAGGGGTATAGTTCCAGGCTTGCTGTATATACCCAGAAATATAGCAAACTCTGG 2956
Qy 227 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuLeuSerIle 246
Db 2957 ACAGGAATAATTATATCTCTCATCTATGTTGGCAACTAACACTGTTACTCTTAGCAATT 3016
Qy 247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266
Db 3017 GTACCCCATCTTGAATAGCAGGAGTTGTTGAATGAAATGTTGCTGGCAACAGCATG 3076
Qy 267 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286
Db 3077 AAGATAGAAGAACTAGAAAGTGTCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTC 3136
Qy 287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306
Db 3137 CGAACCGTTGTTCTTTGACTCAGAGCAGAAAGTTTGAACATATGATGCTCAGAGTTTG 3196
Qy 307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326
Db 3197 CAGGTACCATACAGAACTCTTTGAGGAAAGCAGACATCTTTGGAAATTTACATTTTCTTC 3256
Qy 327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyValaTyrLeuIle 346
Db 3257 ACCCAGGCAATGATGATTTTCTATGCTGGATGTTCCGGTTTGGAGCTTACTTGGTG 3316
Qy 347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366
Db 3317 GCACATAAATCATGAGCTTTGAGGATGTTCTGTGTAGTATTTTTCAGCTGTTGTCTTGGT 3376
Qy 367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386
Db 3377 GCCATGGCCGTGGGCAAGTCAAGTTCTCTGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCT 3436
Qy 387 AlaAlaHisLeuPheAlaLeuLeuLysLysLysProAsnIleAspSerArgSerGlnGlu 406
Db 3437 GGAGCCCAATCATCATGATCATTTGAAAAAACCCCTTTGATTTGACAGCTACAGCAGGAA 3496
Qy 407 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 426
Db 3497 GGCCTAATGCGAACACATTCGAAGAAATGTTCACATTTTGGTGAAGTTGTATTCAACTAT 3556
Qy 427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446
Db 3557 CCCACCCGACCGGACATCCAGTGCTTCAGGAGCTGAGCCCTGGAGGTGAAGAGGCGCAG 3616
Qy 447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 466
Db 3617 ACCTGCTCTGTTGGGCGAGCAGTGGCTGTGGGAAGACACACAGTGTCTCAGCTCTCTGAG 3676
Qy 467 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 486
Db 3677 CGGTTCTACGACCCCTTGGCAGGGGAAAGTGTCTGTGATGGCAAGAAATAAAGCGACTG 3736
Qy 487 AsnValGlnTyrLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506
Db 3737 AATGTTCTAGTGGCTCCGAGACACCTCGGCGCATCGTGTCCAGGAGCCCATCTCTGTTGAC 3796
Qy 507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 526
Db 3797 TGCACATTTGCTGAGAAATTCCTGATGAGACAAACACCGCGGTGTGTACAGGAAGAG 3856
Qy 527 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546
Db 3857 ATCTGAGGGCGCAAGAGGCGCAACATACATGCTTTCATCGACTGCTGCTAATAAA 3916
Qy 547 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 566
```


Db 3917 TATAGCACTAAAGTAGGAGACAAGGAACTCAGCTCTCTGGTGGCCAGAAAACAGCATT 3976
QY 567 AlaileAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 586
Db 3977 GCCATAGCTCGTCCCTTTGGTAGACAGCTCATATTTTGGCTTTGGATGAAGCCAGCTCA 4036
QY 587 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606
Db 4037 GCCTCGGATACAGAAAGTGAAGAGTTGTCCAAAGAGCCCTGGACAAAGCCAGAGAGGC 4096
QY 607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626
Db 4097 CGCACTGCAATGTGATGTCTACCGCTGTCCACCATCCAGAAATCGACACTTAATAGTG 4156
QY 627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 646
Db 4157 GTGTTTCAGAAATGGCAGAGTCAAGGAGCAGTGGCAGCATCAGCAGCTGCTGGCACAGAAA 4216
QY 647 AspileTyrPheLysLeuValAsnAlaGlnSer 657
Db 4217 GGCATCTATTTTCAATGGTCACTGCTCCAGGCT 4249

RESULT 6

US-08-793-610-5
; Sequence 5, Application US/08793610
; Patent No. 5858744
; GENERAL INFORMATION:
; APPLICANT: BAUM, Christopher
; APPLICANT: STOCKING-HARBERS, Carol
; APPLICANT: OSTERTAG, Wolfram
; TITLE OF INVENTION: RETROVIRAL VECTOR HYBRIDS AND THE USE THEREOF
; TITLE OF INVENTION: FOR GENE TRANSFER
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nikaido, Marmelstein, Murray & Oram LLP
; STREET: 655 Fifteenth Street N.W. Suite 330
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/793,610
; FILING DATE: 07-MAR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 44 31 973.8
; FILING DATE: 08-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 195 03 952.1
; FILING DATE: 07-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP95/03175
; FILING DATE: 10-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Berman, Richard J.
; REGISTRATION NUMBER: 39,105
; REFERENCE/DOCKET NUMBER: P1614-7007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)638-5000
; TELEFAX: (202)638-4810
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6505 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA
US-08-793-610-5

Alignment Scores: 1.3e-218 Length: 6505
Pred. No.: 1872.00 Matches: 356
Score: 74.22% Conservative: 142
Best Local Similarity: 53.06% Mismatches: 159
Query Match: 56.33% Indels: 14
DB: 2 Gaps: 4
US-09-873-409-1 (1-659) x US-08-793-610-5 (1-6505)
QY 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
Db 3629 GTCATTTGGAGAAAGAAATCATGATGAACCTCAAGAAAGAAAGGCAATTTCTCTCAA 3688
QY 21 LeuValMetSerGln-----AspIleLysLysAlaAspGluGlnMet 34
Db 3689 CTTGTCAACATGCACACAGCAGGAAATGAAGTTGAATTAGAAAAATGCAGCTGATCAATCC 3748
QY 35 GluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--- 53
Db 3749 AAAAGTGAAATTTGATGCCTTGGAATGTCTCAATATGATTCAAGATCCAGTCTAATAAGA 3808
QY 54 -----LysSerIleLys---SerAspPheIleAspLysAlaGluGluSer 67
Db 3809 AAAAGATCAACTGTAGGAGTGTCCGTGGATCACAAAGCCCAAGACAGAAAGCTTAGTAGCC 3868
QY 68 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 87
Db 3869 AAAGAGGCTCTGGAGAAAGTATACCTCCAGTTCTCTTTGGAGGATTTATGAAGCTAAAT 3928
QY 88 LysProGluTyrPropheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107
Db 3929 TTAAGTGAATGGCCTTATTTTGTGTGTGTATTTTGTGTCATTTATTAATGGAGGCGCTG 3988
QY 108 HisProValPheSerIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAsp 126
Db 3989 CAACCGCATTTTCAATATATTTTCAAGATTTATAGGGGTTTTTACAGAAATTTGATGAT 4048
QY 127 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 146
Db 4049 CCGTGAACCAAAACGACAGATAGTAACTTGTTTTCACATTTGTTTCTAGCCCTTGAATTT 4108
QY 147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 166
Db 4109 ATTTCTTTTATTTACATTTTCTCCTCAGGGTTTTCATTTGGCAAGCTGGAGAGATCCTC 4168
QY 167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 186
Db 4169 ACCAAGCGCTCCGATACATGTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTT 4228
QY 187 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206
Db 4229 GATGACCCCTAAAACACACCACCTGGAGCATTTGACTACAGGCTCGCCCAATGATCTCTCAA 4288
QY 207 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226
Db 4289 GTTAAAGGGGCTATAGGTTCCAGGCTTCTGTAATTTACCCAGAAATATAGCAAAATCTTGG 4348
QY 227 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleuSerIle 246
Db 4349 ACAGAAATAATTTATCTCTTATCTATGTTGGCAACTTAACACTGTTTACTTTAGCAATT 4408
QY 247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266
Db 4409 GTACCCATCATTTGCAATAGCAGGAGTTGTTGAAATGAAATGTTTCTCTGGACAGCACTG 4468
QY 267 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286
Db 4469 AAAGATAAGAAAGAACTAGAGGTTCTGGGAGATCGCTACTGGAAGCAATAGAAAACTTC 4528
QY 287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306
Db 4529 CGAACCGTTGTTTCTTACTCAGGAGCAGAAAGTTTGAACATATATGATGTCAGAGATTG 4588

QY	307	GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe	326
DB	4589	CAGGTACCATTACAGAACTCTTTGGAGGAAGCACACATCTTTGGAAATTACATTTTCTTCTC	4648
QY	327	SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle	346
DB	4649	ACCCAGGCATGATGATTTTTCCTATGCTGGATGTTTCCGGTTTGAGAGCTACTTGGTG	4708
QY	347	GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly	366
DB	4709	GCACATAAACTCATGAGCTTGAGGATGTTCTGTAGTATTTTCAGCTGTTGTCTTTGGT	4768
QY	367	AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly	386
DB	4769	GCCATGGCGTGGGCAAGTCAGTTTCATTTGCTCTGACTATGCCAAAGGCCAAATATCA	4828
QY	387	AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu	406
DB	4829	GCAGGCCCATCATCATGATCATTTGAAAACCCCTTTGATTGACACTACAGCAGCGAA	4888
QY	407	GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPheTyr	426
DB	4889	GGCTTAATGCCGAACACATTGGAAGAAATGTCACATTTTGGTGAAGTTGTATTCAACTAT	4948
QY	427	ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys	446
DB	4949	CCCACCCGACCGGACATCCAGTGCTTCAGGACCTGAGCCCTGGAGGTGAAAGGGCCAG	5008
QY	447	ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln	466
DB	5009	ACCGTGGCTCTGTGGGACAGTAGTGCTGTGGGAAGACAGATGTTGCCAGCTCCTGGAG	5068
QY	467	ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu	486
DB	5069	CGGTTCTACGACCCCTTGGCAGGAAAGTGCTGCTTGATGGCAAGAAATAAAGCGACTG	5128
QY	487	AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn	506
DB	5129	ATGTGTCAGTGGCTCGGACACACTGGGCATGCTGTCCCAGGAGCCCATCTGTGTTGAC	5188
QY	507	CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValProLeuAspGlu	526
DB	5189	TGCAGCATTTGCTGAGAACATTGGCTATGGAGACACAGCCGGTGGTGTCTACAGGAAGAG	5248
QY	527	IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys	546
DB	5249	ATCGTAGGGCAGCAAAAGGAGGCCAACATACATGCTTTCATCGAGTCACTGCCTTAATAAA	5308
QY	547	TyrAsnThrGlnValGlyLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu	566
DB	5309	TATAGCACTAAAGCTAGGAGACAAAGAACTCAGCTCTCTGTGGTGCCAGAAACAACGCATT	5368
QY	567	AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer	586
DB	5369	GCCATAGCTCGTCCCTTGTATAGACAGCTCATATTTTGGTATTGGATGAAGCCACGCTCA	5428
QY	587	AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly	606
DB	5429	GCTCTGGATACAGAAAGTGAAGAGTTGTCCAAGAAAGCCCTGGACAAAGCCAGAGAAAGGC	5488
QY	607	ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal	626
DB	5489	CGACCTTGCATTGTGATGTCTACCCGCTGTCTCCACATCCAGATGTCAGACTTAATAGT	5548
QY	627	ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg	646
DB	5549	GTGTTTCAGAATGGCAGATCAAGGAGCATGGCCGATCAGCAGTCTGCTGGCACAGAA	5608
QY	647	AspIleTyrPheLysLeuValAsnAlaGlnSer	657
DB	5609	GGCATCTATTTTCAATGGTCACTGTCCAGGCT	5641

QY 54 -----lysSerlleLys---SerAspPheIleAspLysAlaGluGluSer 67
Db 3768 AAAAGATCAACTCGTAGGAGTGTCCTGGGATCAACAAGCCCAAGAGCTTATGACC 3827
QY 68 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeu 87
Db 3828 AAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCCTTTGGAGGATTATGAAGCTAAAT 3887
QY 88 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107
Db 3888 TTAACCTGAATGGCTTATTTGTGTGTGTATTTGTGTCATTAATAATGAGGCGCTG 3947
QY 108 HisProValPheSerllePheAlaLysIleIleThrMetPheGlyAsn---AsnAsp 126
Db 3948 CAACAGCATTTTCAATAATATTTTCAAAAGATTATAGGGGTTTTTACAAGAAATTGATGAT 4007
QY 127 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 146
Db 4008 CCTGAAACAAACGACAGAAATAGTAATCTTGTCTTCTACTATTTGTCTTACCCCTTGGAAAT 4067
QY 147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 166
Db 4068 ATTTCTTTTATTACATTTTCTTCCAGGTTTCCACATTTGGCAAGCTGGAGAGATCCTC 4127
QY 167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 186
Db 4128 ACCAAGCGGCTCCGATACATAGTGTGTTTCCGATCCATCTCAGACAGGATGTGAGTTGGTTT 4187
QY 187 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206
Db 4188 GATGACCTTAAACACCACTGGAGCATTTGATCCAGGCTGCCAATGATGCTGCTCAA 4247
QY 207 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226
Db 4248 GTTAAAGGGGCTATAGTTCAGGCTTGCTGTAATTAACCAAGATATAGCAATCTTGGG 4307
QY 227 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetPhePheLeuIleLeuSerIle 246
Db 4308 ACAGGAATAATTATATCTTATCTATGTTGGCAAGTTTGAACATATGTTACTCTTAGCAATT 4367
QY 247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266
Db 4368 GTACCATCATTCGAATAGCAGGAGTTGTGAATGAAATGTAATGTGTGGACAAGCAGCTG 4427
QY 267 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286
Db 4428 AAGATAAGAAAGAACTAGAAAGTGTGGGAAGATCGCTACTGAAGCAATAGAAACTTC 4487
QY 287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306
Db 4488 CGAACCCTGTTGTTCTTTGACTCAGGAGCAGAAGTTTGAACATATGTTATGCTCAGAGTTTG 4547
QY 307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326
Db 4548 CAGGTACCATACAGAACTCTTTGAGAAAGACACATCTTTGGAATTACATTTTCCTTC 4607
QY 327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 346
Db 4608 ACCAGGCAATGATGATTTTCTTATGCTGGATGTTCCGGTTTGGAGCCTACTTGGTG 4667
QY 347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366
Db 4668 GCACATAAACTCATGAGCTTTGAGGATGTTCTGTAGTATTTTTCAGCTGTTCTTGTGT 4727
QY 367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386
Db 4728 GCATGGCCGCGGGAAGTCACTTTCCTTGTGACTATGCCAAAGCCAAATATCA 4787
QY 387 AlaAlaHisLeuPheAlaLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 406
Db 4788 GCAGCCCATCATCATGATCATTTGAAAAACCCCTTTGATTGACAGCTACAGCAGGAA 4847
QY 407 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPheTyr 426

Db 4848 GGCCTAATGCCGAACACATTTGGAAGCAATGTCAATTTGGTGAAGTTGTATTCAACTAT 4907
QY 427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleLeuArgGlyLys 446
Db 4908 CCCACCCGACCGGACATCCAGTGTCTCAGGACTGAGCCTGGAGGTGAAGAGGCCAG 4967
QY 447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValIcInLeuLeuGln 466
Db 4968 ACGCTGGCTCTGTGGGCGAGCAGTGGCTGTGGAAAGACACAGTGGTCCAGCTCTCGGAG 5027
QY 467 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 486
Db 5028 CGGTTCTACGACCCCTTGGCGAGGAAGTGTCTGTGTCGCAAAATAAAGCGACTG 5087
QY 487 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506
Db 5088 AATGTTCTAGTGGCTCCGAGCACCTGGGCATCGTGTCCAGGAGCCCATCTCTGTTGAC 5147
QY 507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValProLeuAspGlu 526
Db 5148 TGCAGCATTTGCTGAGAACATTTGCCCTATGGAGACACAGCCGGTGGTGTACAGGAAGAG 5207
QY 527 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546
Db 5208 ATCGTGAGGGCAGCAAGAGGAGGCCAACATACATGCTTTCATCGAGTCACTGCTTAATAA 5267
QY 547 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 566
Db 5268 TATAGCATTAAGTAGGAGACANAGGAACCTAGCTCTCTGTGGCCAGAACACAGCAATT 5327
QY 567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 586
Db 5328 GCATAGCTCGTGCCCTTGTAGACAGCCTCATATTTTGTCTTGTGATGAAGCCAGTCA 5387
QY 587 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606
Db 5388 GCTCTGGATACAGAAAGTGAAGGTTTCCAAGAAAGCCTGGACAAAGCCAGAGAAGGC 5447
QY 607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626
Db 5448 CGCACTCGATTGTGATTTGCTCACCGCTGTCCACCATCCAGATGCAGACTTAATAGTG 5507
QY 627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 646
Db 5508 GTGTTTCAGATGGCAGATCAAGGAGCATGGCAGCATCAGAGCTGCTGGCACAGAAA 5567
QY 647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657
Db 5568 GGCATCTATTTTCAATGGTCACTGTCAGGCT 5600
RESULT 8
US-08-752-447-1
; Sequence 1, Application US/08752447
; Patent No. 5994088
; GENERAL INFORMATION:
; APPLICANT: Mechethner, Eugene
; APPLICANT: Roninson, Igor B
; TITLE OF INVENTION: Methods and Reagents for Preparing and
; TITLE OF INVENTION: Using Immunological Agents Specific for P-glycoprotein
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESS: McDonnell Boehrnen Hulbert & Berghoff Ltd.
; STREET: 300 South Wacker Drive, Seventh Floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/752.447
FILING DATE: 15-NOV-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: No. 5994088nan, Kevin E
REGISTRATION NUMBER: 35,303
REFERENCE/DOCKET NUMBER: 95,1121
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-913-0001
TELEFAX: 312-913-9808
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4669 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..424
FEATURE:
NAME/KEY: CDS
LOCATION: 425..4264
FEATURE:
NAME/KEY: 3'UTR
LOCATION: 4265..4669
US-08-752-447-1

Alignment Scores:
Pred. No.: 8,88e-217 Length: 4669
Score: 1855.00 Matches: 354
Percent Similarity: 73.47% Conservative: 139
Best Local Similarity: 52.76% Mismatches: 164
Query Match: 55.82% Indels: 14
DB: 2 Gaps: 4

US-09-873-409-1 (1-659) x US-08-752-447-1 (1-4669)

Qy 1 MetLeuAlaGluValGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrSer 20
Db 2237 GTCACTTGGGAGAAAGGAATCATGATGAACCTCAAGAAAGAGGAGGCAATTTACTTCAA 2296
Qy 21 LeuValMetSerGln-----AspIleLysLysAlaAspGluGlnMet 34
Db 2297 CTTGTTCACATCGACACAGCAGCAAGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCC 2356
Qy 35 GluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--- 53
Db 2357 AAAAGTGAATGATGCCTTGGAAATGTCTTCAAAATGATTCAAGATCCAGTCTAATAAGA 2416
Qy 54 -----LysSerIleLysSerAspPheIleAspLysAlaGluSerThr 68
Db 2417 AAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAGCCCAACAGAAAGCTTAGTACC 2476
Qy 69 ---GlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 87
Db 2477 AAAGAGGCTCTGGATGAAGTATACCTCCAGTTCTTTGGAGGATATAGAGCTAAT 2536
Qy 88 LysProGluTyrProPheValValLeuGluThrLeuAlaSerValLeuAsnGlyThrVal 107
Db 2537 TTAACCTGAATGCGCTATTTTGTGTGTGTATTGTGTGCCATTATATAATGGAGCGCTG 2596
Qy 108 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAsp 126
Db 2597 CAACGAGCATTTTGCATAATAATATTTCAAAAGATTATAGGGGTTTTCACAGAAATTGATGAT 2656
Qy 127 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 146
Db 2657 CCGTGAACCAAAACGACAGCAATAGTAACTTGTCTTCACTATTGTTTCTAGCCCTTGAAT 2716
Qy 147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 166
Db 147 TGCAGCATTTGCTGAGAACATTTGCCTATGGAGACAAACAGCCGGGTGTGTGTACAGGAAGAG 3856

Db 2717 ATTTCTTTTATTACATTTTCCCTTCAGGGTTTTCACATTTGGCAAAGCTGGAGATCCTC 2776
Qy 167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPhe 186
Db 2777 ACCAAGCGGCTCCGATACATGTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTT 2836
Qy 187 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206
Db 2837 CATGACCCCTAAAAACACCACCTGGAGCATTTGACTACAGGCTCGCAATGATGCTGCTCAA 2896
Qy 207 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226
Db 2897 GTTAAAGGGGCTATAGTTCCAGGCTTGTCTGTAATACCCAGATATATAGCAAAATCTTGG 2956
Qy 227 LeuSerValIleIleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIle 246
Db 2957 ACAGGAATATATATCTCTCATCTATGTTGGCACTAAACACTGTTACTCTTAGCAATT 3016
Qy 247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266
Db 3017 GTACCCATCATTTGCAATAGCAGGAGTTGTTGAAATGAAATGTTTGTCTGGCAAGCCTG 3076
Qy 267 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286
Db 3077 AAAGATAGAAGAACTAGAAAGGTCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTC 3136
Qy 287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306
Db 3137 CGAACCGTTGTTCTTTGACTCAGGAGCAGAAAGTTTGAACATATGATGCTCAGAGTTTG 3196
Qy 307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326
Db 3197 CAGGTACCATACAGAAACTCTTTGAGGAAGACACACATCTTTGGAAATTTACATTTTCTTCT 3256
Qy 327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyValThrIleuIle 346
Db 3257 ACCGAGGCAATGATGATTTTCTTATCTGGATGTTTCCGGTTTGGAGCCCTACTTGTG 3316
Qy 347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366
Db 3317 GCACATAAATCATGAGCTTTGAGGATGTTCTGTAGTATTTTTCAGCTGTTGTCTTTGTT 3376
Qy 367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386
Db 3377 GCCATGGCGTGGGCAAGTCAGTTCTTCTGCTGCTGCTATGCCAAAGCCCAAAATATCA 3436
Qy 387 AlaAlaHisLeuPheAlaLeuLeuLysLysProAsnIleAspSerArgSerGlnGlu 406
Db 3437 GCAGCCCATCATCATGATCATTTGAAAAAACCCCTTTGATTTGACAGCTACAGCAGGAA 3496
Qy 407 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPheTyr 426
Db 3497 GGCCTAATGCCGACACATTTGGAAGAAATGTCACATTTTGTGAAGTTGTATTCAACTAT 3556
Qy 427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446
Db 3557 CCCACCCGACCGACATCCAGTGTCTCAGGACCTGAGCCTGGAGGTGAAGAGGCCAG 3616
Qy 447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuGln 466
Db 3617 ACGTGGCTCTGGTGGGAGCAGTGGCTGTGGGAAGACACAGTGTCCAGCTCTCTGGAG 3676
Qy 467 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 486
Db 3677 CGTTTCTAGACCCCTTGGCAGGGAAGTGTGTTGATGGCAAGAAATAAAGCGACTG 3736
Qy 487 AsnValGlnTyrLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506
Db 3737 AATGTTCACTGGCTCCGAGCACACCTGGGCATCGTGTCCAGAGGCCCATCTCTGTTGAC 3796
Qy 507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 526
Db 3797 TGCAGCATTTGCTGAGAACATTTGCCTATGGAGACAAACAGCCGGGTGTGTGTACAGGAAGAG 3856

QY	527	IleIysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys	546
Db	3857	ATCGTAGGGCAGCAAGAGGGCCAAACATACATGCTTCCTCATCGAGTCACCTGCCATAATAAA	3916
QY	547	TyrAsnThrClnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu	566
Db	3917	TATAGCACTAAAGTAGGAGACAAGAACTCAGCTCTCTGGTGGCCAGAAAACAACGCAT	3976
QY	567	AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer	586
Db	3977	GCATAGCTCGTCGCCTTGTTAGACAGCCTCATATTTCCTTTGGATGAAGCCACGCTCA	4036
QY	587	AlaLeuAspAsnSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly	606
Db	4037	GCTCTGGATACAGAAAGTGAAGAAGTTGTCCAAGAGCCCTGGACAAGCCAGAGAAGCC	4096
QY	607	ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal	626
Db	4097	CGCACCTGCATTGTGATTGCTCACCGCTGTCCACCATCCAGATCCAGACTTAAATAGTG	4156
QY	627	ValLeuHisGlnGlyLysLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg	646
Db	4157	GTGTTTCAGAATGGCAGAGTCAAGGAGCATGGCAGCATCAGCAGCTGCTGGCACAGAAA	4216
QY	647	AspIleTyrPheLysLeuValAsnAlaGlnSer	657
Db	4217	GGCATCTATTTTTCAATGGTCAGTGCCAGGCT	4249

RESULT 9

US-09-316-167-1
; Sequence 1, Application US/09316167
; Patent No. 6365357
; GENERAL INFORMATION:
; APPLICANT: Mechetner, Eugene
; APPLICANT: Roninson, Igor B
; TITLE OF INVENTION: Methods and Reagents for Preparing and
; TITLE OF INVENTION: Using Immunological Agents Specific for P-glycoprotein
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boenhen Hulbert & Berghoff Ltd.
; STREET: 300 South Wacker Drive, Seventh Floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/316,167
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/752,447
; FILING DATE: 15-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6365357nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 95,1121
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-9808
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4669 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:


```

Db 4097 CGCACCTGCATGTGTCTCACCCTGTCACCATCCAGATGCAGACTTAATAGTG 4155
Qy 627 ValLeuHisAsnGlyLysLeuGlnGlyThrHisGlnLeuLeuArgAsnArg 646
Db 4157 GTGTTTCAGAATGGCAGAGTCAAGGAGCATGGCAGCATCAGCAGCTGCTGGCAGAAA 4216
Qy 647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657
Db 4217 GGCATCTATTTTTCATATGGTCAGTCCAGGCT 4249

RESULT 10
US-09-120-513-1
; Sequence 1, Application US/09120513
; Patent No. 6025160
; GENERAL INFORMATION:
; APPLICANT: Brun, Kimberly
; APPLICANT: Chenery, Richard
; APPLICANT: Ellens, Harna
; APPLICANT: Field, John
; APPLICANT: Yue, Lin
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE
; TITLE OF INVENTION: SEQUENCES ENCODING RAT MDR1B2 AND
; TITLE OF INVENTION: SCREENING METHODS THEREOF
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY:
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09120,513
; FILING DATE: 22-JUL-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: King, William T.
; REGISTRATION NUMBER: 30,954
; REFERENCE/DOCKET NUMBER: GP50008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5015
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4233 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
US-09-120-513-1

Alignment Scores:
Pred. No.: 3,54e-216 Length: 4233
Score: 1849.50 Matches: 358
Percent Similarity: 73.13% Conservative: 132
Best Local Similarity: 53.41% Mismatches: 163
Query Match: 55.66% Indels: 17
DB: 3 Gaps: 5

US-09-873-409-1 (1-659) x US-09-120-513-1 (1-4233)
Qy 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
Db 1832 GTCATTGCGAGCAAGAAATCATGAAGAGCTCATGAAAGAGGAGCATTTACTTCAA 1891
Qy 21 LeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSer 40

```


Qy	407	GlyLysIlePProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr	426
Db	3497	GGCCTAATGCCGAACACATTTGGAAGGAAATGTCATCATTTGGTGAAGTTGATTCAACTAT	3556
Qy	427	ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys	446
Db	3557	CCACCCGACCGGACATCCAGCTTTCAGGACATGAGCCCTGGAGGTGAAGAAGGGCCAG	3616
Qy	447	ThrValAlaPheValGlySerGlyCysGlyLysSerThrSerValGlnLeuLeuGln	466
Db	3617	ACCTGGCTCTGTGGGACAGCTGGCTGTGGAGAGACACAGTGGTCCAGCTCTCTGGAG	3676
Qy	467	ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu	486
Db	3677	CGGTTCTACGACCCCTTGGCAGGAAAGTGCCTTGATGTGCAAGAAATAAAGCGCACTG	3736
Qy	487	AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn	506
Db	3737	ANTGTTCAAGTGGCTCCGACGACACCTGGGCATCTGTGTCCTCAGGAGCCCATCTCTTTGAC	3796
Qy	507	CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu	526
Db	3797	TGCAGCATTCGTGAGAACATTCGCTATGGACACACACGCCGGTGTCTCACAGGAAGAG	3856
Qy	527	IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys	546
Db	3857	ATCGTGAGGCGACAAAGGAGGCCAACATACATGCTTCATCGAGTCACCTGCCTAATAAA	3916
Qy	547	TyrAsnThrClnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu	566
Db	3917	TATAGCACTAAAGCTAGGAGACAAAGGAACCTCAGCTCTCTGGTGGCCAGAAACACGCAT	3976
Qy	567	AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer	586
Db	3977	GCCATAGCTCGTCCCTTGTGTAGACAGCCTCATATTTGCTTTTGGATGAAGCCACGCTCA	4036
Qy	587	AlaLeuAspAsnAspSerGlnLysValValGlnHisAlaLeuAspLysAlaArgThrGly	606
Db	4037	GCTCTGGATACAGAAAGTGAAAGGTTGTCCAAAGAGCCCTGGACAAACCCAGAGAAAGGC	4096
Qy	607	ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal	626
Db	4097	CGCACCTGCATTCGATTGCTCACCAGCTGTCCACCATCCAGAATGCAGACTTAATAGTG	4156
Qy	627	ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnLeuLeuArgAsnArg	646
Db	4157	GTGTTTCAGAATGGCAGAGTCAGACAGCATGGCAGCATCAGCAGCTCTGGCACAGAAA	4216
Qy	647	AspIleTyrPheLysLeuValAsnAlaGln	656
Db	4217	GGCATCTATTTTCAATGGTCAAGTCCAG	4246
RESULT 13			
US-08-996-545-1			
; Sequence 1, Application US/08996545			
; Patent No. 5928898			
; GENERAL INFORMATION:			
; APPLICANT: Skatrud, Paul L.			
; APPLICANT: de Waard, Maarten A.			
; APPLICANT: Peery, Robert B.			
; APPLICANT: Andrade, Alan C.			
; TITLE OF INVENTION: Multiple Drug Resistance Gene at/d of			
; TITLE OF INVENTION: Aspergillus nidulans			
; NUMBER OF SEQUENCES: 3			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: Eli Lilly and Company			
; STREET: Lilly Corporate Center			
; CITY: Indianapolis			
; STATE: Indiana			
; COUNTRY: U.S.			
; ZIP: 46285			
; COMPUTER READABLE FORM:			

Db 2434 ATGTTCTTCGTGGTGGTATCATTCAGTTTATACGACGATCAACCAATGGTGTGCAATTT 2493
Qy 160 GlyArgAlaGlyGluLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeu 179
Db 2494 GCGGTATGCTCCGAGAGACTTATTCGTCGCGGAGAGCACTGCCTTTCCGAGCATCTC 2553
Qy 180 TyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrIle 199
Db 2554 CGTCAAGACATTCCTTCTTGACAGAGAGAGAAATAGACCGCGCTCTGACCTCTTC 2613
Qy 200 LeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThr 219
Db 2614 CTGTCCACGAGAGAGCATCTCTCCGTTGTAGCGGTGTGACTCTAGCAGCATCTTG 2673
Qy 220 GlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMet 239
Db 2674 ATGACCTCCACGACCCTAGGAGCGCTATCATTTATTCCTGCGCATGGTGGAAATG 2733
Qy 240 ThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAla 259
Db 2734 GCCTTAGTTTGTATCTCGTTGTGCGGGTCTCTCTGCGCATGGCTTCTACCGATTCTAT 2793
Qy 260 AlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAla 279
Db 2794 ATGCTAGCCAGTTTCAATCAGCTCTCAAGCTGTCTTATGAGGATCTGCAAACTTTGCT 2853
Qy 280 ThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGlu 299
Db 2854 TCGGAGGCTACATCTATCCGACAGTTTGGCTATTAAACCGCGGAAAGGATGCTCG 2913
Qy 300 GlnMetTyrGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIle 319
Db 2914 GAGATTATACCATGCCAGCTTGACGACAGGCGAGCAGCATCTAATCTCTGCTTTGAGG 2973
Qy 320 IleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPhe 339
Db 2974 TCATCCTGTATATGCGTCGTCGAGGACACTGTGTTTCTTCGCGTTGCGTCGCGTTT 3033
Qy 340 ArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleVal 359
Db 3034 TGCTACGGAGGGACACTTCTTGCTGTCACCGAGTATGACATTTTCGCTCTTGTGTTGT 3093
Qy 360 PheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGlu 379
Db 3094 TTCTCCGAGATTCTCTTGTGTCATCCTCGCGGACCGCTCTTTTCTTTCACCCAGAC 3153
Qy 380 TyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsn 399
Db 3154 ATGGCAAGCGCAAGAAATGCGCGCGCGAATTCGACGACTGTTCGACCGAAGCCACAA 3213
Qy 400 IleAspSerArgSerGlnGlyLysLysProAspThrCysGluGlyAsnLeuGluPhe 419
Db 3214 ATTGATACTGTTCTCAAGAGGCGGAGAGCTCGAACGGTGAAGGTGAATCGAATTT 3273
Qy 420 ArgGluValSerPheTyrProCysArgProAspValPheIleLeuArgGlyLeuSer 439
Db 3274 AGGAACGTCGACTTCAGATACCCGACCGCCAGAACAGCCTGTCTCGCGCGCTTGGAC 3333
Qy 440 LeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSer 459
Db 3334 CTGACCGTGAAGCCTGGACAATATGTGTGCTGTTCGACCCAGCGGTGTGGCAAGAT 3393
Qy 460 ThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAsp 479
Db 3394 ACCACCATTCATGCTTGAGCGCTTTTACGATGCGATTGCGGGTCCATCTCTTGTGAT 3453
Qy 480 GlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValPro 499
Db 3454 GGGAGGACATAGTAACCTAAATATCACTCTCCTACCGAGCTTTCTGTCACTGTGTGAC 3513
Qy 500 GlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSer 519
Db 3514 CAGGACCGGACACTGTACCGGCGCACCATCAAGAAACATCTTACTTGTATTGTGAA 3573

Qy 520 ArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPhe 539
Db 3574 CATGACGTACCGGAAGAATTTCTTGATTAAAGGCTTCAAGGACGCTAATATCTACGACTTC 3633
Qy 540 IleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyValAlaGlnLeuSer 559
Db 3634 ATCATGTGCTCCCGAGGCTTTAATACAGTTGTGGCAGCAAGGAGGAGCATGTGTGCT 3693
Qy 560 GlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeu 579
Db 3694 GCGGCGCAAAAGCAACGTGTGCCATTGCCGAGGCCCTTCTTCGGGATCCCAAAATCCTT 3753
Qy 580 LeuLeuAspGluAlaThrSerAlaLeuAspAspSerGluLysValValGlnHisAla 599
Db 3754 CTTCCTCATGAAGCGACGCTCAGCCCTCAGTCCGAGTCAGAAAGGTCGTCCAGCGGCT 3813
Qy 600 LeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIle 619
Db 3814 TTGGATGCGCTGCCGAGCGCGAACCAATCGCGTTGACACCGACTCAGCAGCATT 3873
Qy 620 GlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHis 639
Db 3874 CAAAAGGCGGAGCTTATCTATGTTTTCGACCAAGCAAGATCGTCGAAAGCGGAACGAC 3933
Qy 640 GlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerVal 658
Db 3934 AGCGAACTGGTCCAGAAAAAGGCGCGGTACTACGAGCTGGTCAACTGTCAGAGCTTG 3990

RESULT 14
US-08-996-545-3
; Sequence 3, Application US/08996545
; Patent No. 5928898
; GENERAL INFORMATION:
; APPLICANT: Skatrud, Paul L.
; APPLICANT: de Waard, Maarten A.
; APPLICANT: Feery, Robert B.
; APPLICANT: Andrade, Alan C.
; TITLE OF INVENTION: Multiple Drug Resistance Gene atRD of
; TITLE OF INVENTION: Aspergillus Nidulans
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: U.S.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/996,545
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Webster, Thomas D.
; REGISTRATION NUMBER: 39,872
; REFERENCE/DOCKET NUMBER: X-11766
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3334
; TELEFAX: 317-276-2763
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4002 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO

US-08-996-545-3

Alignment Scores:

Pred. No.:	4,67e-143	Length:	4002
Score:	1254.00	Matches:	274
Percent Similarity:	58.47%	Conservative:	123
Best Local Similarity:	40.35%	Mismatches:	260
Query Match:	37.74%	Indels:	22
DB:	2	Gaps:	6

US-09-873-409-1 (1-659) x US-08-996-545-3 (1-4002)

Qy 2 LeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeu 21
Db 1954 AUUGCUGAAGCAACACCAUGAAUUGGUGAGCCGCGAGGCGCUUUAUGCAACUU 2013

Qy 22 ValMetSerGlnAspIleLysAlaAspGluGlnMetGluSerMetThrTyrSerThr 41
Db 2014 GUGGAGGCGUACAAGUAACAGAAAGGAGGUGAGCCUUGGAGGAGCCCGACGCU 2073

Qy 42 GluAArgLysThrAsnSer-----LeuProLeuHisSerValLysSerIleLysSerAsp 59
Db 2074 GAGGAUCUCACAGAUAGCAUAUUGCCAAAUCAAAACUGCGUACGCGCAUCAUCGGAU 2133

Qy 60 Phe-----IleAspLys-----AlaGluGluSerThrGlnSerLys 71
Db 2134 CUCGACGGAAACCCACACCAUUGACCGCACGGGACCCACCAAGUCUGUUCACGCGG 2193

Qy 72 GluIleSer-----LeuProGluValSerLeuLysIleLysIleLys 85
Db 2194 AUUCUUUCAAAGAACCCCGAAACAAACUCCGAAUAUCUAUAUUGGACGCGUCGCAAA 2253

Qy 86 -----LeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSer 101
Db 2254 UUGUUGUUCUUCUACACCGCCGUAUCCGUAUCGCUACGCUUUGUUCUUCUA 2313

Qy 102 ValLeuAsnGlyThrValHisProValPheSerIlePheAlaLysIleThrMet 121
Db 2314 GUGUAGUGGUGGUGGCGCAACCCACGCAAGCAGUGCUAUUUGCUAAAGCCACACACA 2373

Qy 122 PheGly-----AsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMet 139
Db 2374 CUCUGCUCCCGAGAAUACAUAUAGCAAGCUUCGACAUUGCGGAUUCUGGUCAUUG 2433

Qy 140 IlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyr 159
Db 2434 AUGUUCUUGUGUGUUAUCAUUCAGUUUAUUCACGACGACCAACCAUUGUGUGCAUUU 2493

Qy 160 GlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeu 179
Db 2494 GCCGUAGUCCGAGAGACUUAUUGCGCGGAGAGCACUGCCUUGCGGACGAUACUC 2553

Qy 180 TyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyClyLeuThrThrIle 199
Db 2554 CGUCAAGACAUUGCUUCUUGACAGGAAGAAUAGCACCAGCGGCGUCUGACCUUUC 2613

Qy 200 LeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThr 219
Db 2614 CUUGCCAGCAGCAGCAAGCAUUCUCCGUGUAGCGGUGAGUAGCUUAGCAGCAUUG 2673

Qy 220 GlnAspAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMet 239
Db 2674 AUGACCCUCCAGCACCCUAGGAGCGCUAUCUAUUGCCUUGCGGAUUGGGGAAUUG 2733

Qy 240 ThrPheLeuLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAla 259
Db 2734 GCCUAGUUGUUAUCUGGUGUGCGGUGUUCUCCGGCAGCGGUGUUCUACCGGAUUCUAU 2793

Qy 260 AlaMetThrGlyPheAlaAsnLysAspLysGlnLeuLysHisAlaGlyLysIleAla 279
Db 2794 AUGCUAGCCAGUUUCAAUCGCUACUCCUAGCUUUGUUAUUGAGGGAUUCUGCAACUUGCU 2853

Qy 280 ThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGlu 299

Db 2854 UGCAGGCUACAUCGCUUAUCCGACAGUUGCGUCAUUAACCCGGAAAGGGAUGUUCUGG 2913

Qy 300 GlnMetTyrGluGluMetLeuGlnThrHisArgAsnThrSerLysLysAlaGlnIle 319
Db 2914 GAGAUUAACCAUGCCAGCUUGACGACACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2973

Qy 320 IleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPhe 339
Db 2974 UCAUCCUUGUUAUUGCGUGCGCAGGACUUGUUAUUGCGUGUUGCGUGCGUGUUGU 3033

Qy 340 ArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleVal 359
Db 3034 UGUACGAGGAGGACACUUCUUGGUCACACGAGUAUGACAUUUUCCGCUUCUUGUUGU 3093

Qy 360 PheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGlu 379
Db 3094 UUCUCCGAGAUUCUUCUUGGUGCUCAUCCGCGGCGCAGGAGGAGGAGGAGGAGGAGGAG 3153

Qy 380 TyrSerLysAlaLysSerGlyAlaHisLeuPheAlaLeuLeuLysLysProAsn 399
Db 3154 AUGGCAAGGCGAAGAUUGCGCGCCGAAUUCGACGACUGUUGCGAAGGAGGAGGAGGAGGAG 3213

Qy 400 IleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPhe 419
Db 3214 AUUGAAUACUGGUCUAGAGGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3273

Qy 420 ArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSer 439
Db 3274 AGGAAGUGGACUACAGAUACCCGACCCGACGACGAGGAGGAGGAGGAGGAGGAGGAGGAG 3333

Qy 440 LeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSer 459
Db 3334 CUGACCGUGAAGCUGGACAAUAUUGCGCUGUUGCGGACCCGACGAGGAGGAGGAGGAGGAG 3393

Qy 460 ThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAsp 479
Db 3394 ACCACAUUGCAUUGCUUGAGCGCUUUAUACGAGUUGCGGCGGAGGAGGAGGAGGAGGAG 3453

Qy 480 GlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValPro 499
Db 3454 GGAAGGACAUAGUA 3513

Qy 500 GlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSer 519
Db 3514 CAGGAGCGACACUGUACCGGCGACCAUACAAAGGAAACAUUUAUUGGUAUUGGUAUUG 3573

Qy 520 ArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPhe 539
Db 3574 GAUGACGUACCGGAAGAUUCUGAUUAAGGCUUGCAAGGAGGAGGAGGAGGAGGAGGAGGAG 3633

Qy 540 IleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyValaGlnLeuSer 559
Db 3634 AUCAUGCGUCCCGGAGGCGCUUUAUAUAUAUAUAUAUAUAUAUAUAUAUAUAUAUAUAUA 3693

Qy 560 GlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeu 579
Db 3694 GCGGCGCAAAAGCAACUGUGGCAUUGCCGAGGCGCUUUAUUGGAGGAGGAGGAGGAGGAGGAG 3753

Qy 580 LeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAla 599
Db 3754 CUUCUGAUGAAGCGACGCGCUCGACUCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3813

Qy 600 LeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIle 619
Db 3814 UUGGAGCGCGUCCCGGAGGCGCAACCAUUGCGCGUUGGACACCCGACGACGAGGAGGAGGAG 3873

Qy 620 GlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHis 639
Db 3874 CAAAAGGCGGAGCUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 3933

Qy 640 GlnGluLeuLeuArgAsnArgIleTyrPheLysLeuValAsnAlaGlnSerVal 658
Db 640 GlnGluLeuLeuArgAsnArgIleTyrPheLysLeuValAsnAlaGlnSerVal 658

Db 3934 AGCAACUGGUCCAGAAAAGGCGCGGUACUACGAGCUGGUCAACUUGCAGACCUUG 3990

RESULT 15

US-09-328-320-1

; Sequence 1, Application US/09328320

; Patent No. 6228615

; GENERAL INFORMATION:

; APPLICANT: Skatrud, Paul L.

; APPLICANT: de Waard, Maarten A.

; APPLICANT: Peery, Robert B.

; APPLICANT: Andrade, Alan C.

; TITLE OF INVENTION: Multiple Drug Resistance Gene atpD of

; TITLE OF INVENTION: Aspergillus nidulans

; NUMBER OF SEQUENCES: 3

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Eli Lilly and Company

; CITY: Indianapolis

; STATE: Indiana

; COUNTRY: U.S.

; ZIP: 46285

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/328,320

; FILING DATE:

; CLASSIFICATION:

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: 08/996,545

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Webster, Thomas D.

; REGISTRATION NUMBER: 39,972

; REFERENCE/DOCKET NUMBER: X-11766

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 317-276-3334

; TELEFAX: 317-276-2763

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 4002 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 1..4002

US-09-328-320-1

Alignment Scores:

Pred. No.: 4,67e-143 Length: 4002

Score: 1254.00 Matches: 274

Percent Similarity: 58.47% Conservative: 123

Best Local Similarity: 40.35% Mismatches: 260

Query Match: 37.74% Indels: 22

DB: 4 Gaps: 6

US-09-873-409-1 (1-659) x US-09-328-320-1 (1-4002)

Qy 2 LeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeu 21

Db 1954 ATTCCTGAACAGAACTCACGATGAATTTGGTTGACCGCGGAGGCGCTTATCGCAAACTT 2013

Qy 22 ValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThr 41

Db 2014 GTGGAGGCTCAACGATATCAATGAACAGAGGAAGCTGACGCTTGGAGGACCGCAGCT 2073

Qy 42 GluArgLysThrAsnSer-----LeuProLeuHisSerValLysSerIleLysSerAsp 59

Db 2074 GAGGATCTCAGCAATGCAGATATTGCCAAATCAAAATCGCTCAAGCGCATCATCCGAT 2133

Qy 60 Phe-----IleAspLys-----AlaGluGluSerThrGlnSerLys 71

Db 2134 CTCGAGGAAACCCACAAACCATTCGCGCAGGCGCACCAAGCTGTGTTCCAGCGG 2193

Qy 72 GluIleSer-----LeuProGluValSerLeuLeuLysIleLeuLys 85

Db 2194 ATTCCTTTCTAAAGACCCCGGAAACAACTCCGAAATACTCATTATGAGCGCTGTCTCAA 2253

Qy 86-----LeuAsnLysProGluTrpProPheValLeuGlyThrLeuAlaSer 101

Db 2254 TTTGTGTCTTCCTTCAACGCGCTGAAATCCGTAACATCATCGTCTCATCGTCTTCTCTCA 2313

Qy 102 ValLeuAsnGlyThrValHisProValPheSerIleLeuPheAlaLysIleThrMet 121

Db 2314 GGTGTAGCTGTGTGGCCCAACCCAGCAAGCAGTGTATATGTCTAAAGCCATCAGCAC 2373

Qy 122 PheGly-----AsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMet 139

Db 2374 CTCTCGCTCCCGAATCAATATAGCAAGCTTCGACATGATGCGGATTTCTGGTCAATTG 2433

Qy 140 IlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyr 159

Db 2434 ATGTTCTTCGTGGTGTGATCATTCAGTTTATCAGCAGTCAACCAATGGTGTCTCATTT 2493

Qy 160 GlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeu 179

Db 2494 GCCGTATGCTCGAGAGACTTATTCGTCGCGGAGAACACTGCGCTTCGGAGATATC 2553

Qy 180 TyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIle 199

Db 2554 CGTCAAGACATTTCTTTTGACAGGAGAGATAGCACCGCGCTCTGACCTCTTTC 2613

Qy 200 LeuAlaIleAspIleAlaGlnIleGlnAlaThrGlySerArgIleGlyValLeuThr 219

Db 2614 CTGTCCAGCAGAGCAAGCATCTCTCCGCTGTAGCGGTGTGACTCTAGGACGATCTTG 2673

Qy 220 GlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyrGluMet 239

Db 2674 ATGACCTCCAGCCCTAGGAGCGCTATCATTTATGCCCTGGCGGATTTGGTGGAAATTTG 2733

Qy 240 ThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAla 259

Db 2734 GCCTTAGTTTGTATCTCGGTTGTGCGGTTCTCTCGGCATGCGGTTTCTACCGATCTAT 2793

Qy 260 AlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAla 279

Db 2794 ATGTAGCCCGAGTTTCAATCAGCTCCAAAGCTTGCTTATGAGGATCTGCAAACTTTGCT 2853

Qy 280 ThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGlu 299

Db 2854 TCGAGGCTACATCGTCTATCCGACATTCGCTATTAACCCGGAAGGAGTGTCTGG 2913

Qy 300 GlnMetTyrGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIle 319

Db 2914 GAGATTTACCATGCCAGCTTGACGCAACAGGAGGACAGCTAAATCTCTGCTCTTGGG 2973

Qy 320 IleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPhe 339

Db 2974 TCATCCCTGTTTATATGCTGCTCGCAGGCACTTGTCTTCTGCGTGTGCGCTCGGGTTT 3033

Qy 340 ArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleVal 359

Db 3034 TGGTACGAGGAGGACACTTCTTTGGTCACCGACAGATATGACATTTCCCGCTCTTTTGTGT 3093

Qy 360 PheThrAlaIleAlaTyrGlyValaMetAlaIleGlyLysThrLeuValLeuAlaProGlu 379

Db 3094 TTCTCCGAGATTCCTTTGGTGTCTCAATCCGCGGCGACCGCTTTTCTTCTTGCACGAC 3153

Qy 380 TyrSerLysAlaLysSerGlyAlaHisLeuPheAlaLeuLeuGluLysLysProAsn 399

Db 3154 ATGGGCAAGCGAAGATGCGCGCGCGAATTCGACGACTGTCGACCGGAAGCCACAA 3213
 Qy 400 IleAspSerArgSerGlnGluGlyLysProAspThrCysGluGlyAsnLeuGluPhe 419
 Db 3214 ATTGATAACTGGTCTGAAGAGGCGGAGAGCTCGAACCGGTGGAAGGTGAATCGAATTT 3273
 Qy 420 ArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSer 439
 Db 3274 AGGAACGTGCACCTTCAGATACCGCCGCGCAGAACACGCTGCTCGCGCGCTTGGAC 3333
 Qy 440 LeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSer 459
 Db 3334 CTGACCGTGAAGCCTGGACAATATGTCGGCTTTCGACCCGCGGTGTGGCAAGAGT 3393
 Qy 460 ThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAsp 479
 Db 3394 ACCACCATTCATTCCTTGAGCGCTTTTACGATGCGATTGCGGGTCCATCCTTGTTCGAT 3453
 Qy 480 GlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValPro 499
 Db 3454 GGGAAAGGACATAAGTAATACTAAATATCACTCTACCGCAGCTTCTGTCTACCTGTGTACG 3513
 Qy 500 GlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSer 519
 Db 3514 CAGGAGCCGACACTGTATCCAGGGCACCATCAAGGAAACATCTTACTTGGTATTGTTCGAA 3573
 Qy 520 ArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPhe 539
 Db 3574 GATGACGTACCGGAAGAATTCTTGATTAGGCTTCCAAGGACGCTAATATCTACGACTTC 3633
 Qy 540 IleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSer 559
 Db 3634 ATCATGTCTCCCGAGGGCTTTAATACAGTTGTTGGCAGCAAGGAGGCATGTTGTCT 3693
 Qy 560 GlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeu 579
 Db 3694 GCGGCCCAAAAGCAACGTTGGCCATTGCCGAGCCCTTCTTCGGGATCCCAAAATCCTT 3753
 Qy 580 LeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAla 599
 Db 3754 CTTCTCGATGAAGCGACGTACGCCCTCGACTCCGAGTCAGAAAAGGTCTGTCAGCGGCT 3813
 Qy 600 LeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIle 619
 Db 3814 TTGGATGCGCTGCCGCGCGGCGAACCAATCCGCGTTGCACACCGACTCAGCACGATT 3873
 Qy 620 GlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHis 639
 Db 3874 CAAAAGGCGGACGTTTATCTATGTTTCGACCAAGGCAAGATCGTCGAAAGCGGAACGCAC 3933
 Qy 640 GlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAlaGlnSerVal 658
 Db 3934 ACGGAATGGTCCAGAAAAGGCGCGGTACTACGAGCTGGTCACTTGGCAGAGCTTG 3990

Search completed: March 31, 2003, 15:20:38
 Job time : 192.507 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 31, 2003, 13:52:37 ; Search time 47.683 Seconds
(without alignments)
2337.888 Million cell updates/sec

Title: US-09-873-409-3

Perfect score: 659

Sequence: 1 MVDENDIRALNVRHYRDHIG.....SESKSAVQAALKEKTPRYSF 131

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 574371 seqs, 425486471 residues

Total number of hits satisfying chosen parameters: 1148742

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -DEV=xlh

-MODE=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPTO spooll/US09873409/runat_27032003_115424_19290/app query fasta_1.7544
-DB=Published Applications NA -QWTF=fastcap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-NAXLEN=2000000000 -USPR=US09873409 @CGN 1.1 960 @runat_27032003_115424_19290
-NCPU=6 -ICPU=3 -NO XLPXY -NO MWAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:
1: /cgn2_6/prodata/2/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/prodata/2/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/prodata/2/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/prodata/2/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/prodata/2/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/prodata/2/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/prodata/2/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/prodata/2/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/prodata/2/pubpna/US09_NEW_PUB.seq:
10: /cgn2_6/prodata/2/pubpna/US09_PUBCOMB.seq:
11: /cgn2_6/prodata/2/pubpna/US10_NEW_PUB.seq:
12: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:
13: /cgn2_6/prodata/2/pubpna/US60_NEW_PUB.seq:
14: /cgn2_6/prodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	659	100.0	1175	10	US-09-873-409-11 Sequence 11, Appl
2	659	100.0	1940	10	US-09-873-409-16 Sequence 16, Appl
3	659	100.0	2021	10	US-09-873-409-15 Sequence 15, Appl
4	619	93.9	2856	10	US-09-873-409-10 Sequence 10, Appl

5	619	93.9	3177	10	US-09-873-409-12 Sequence 12, Appl
6	619	93.9	3621	10	US-09-873-409-14 Sequence 14, Appl
7	619	93.9	3702	10	US-09-873-409-13 Sequence 13, Appl
8	468	71.0	4189	10	US-09-866-866A-3 Sequence 5, Appl
9	467	70.9	4254	10	US-09-917-800A-1424 Sequence 1424, Ap
10	467	70.9	4369	10	US-09-769-097-1 Sequence 1, Appl
11	467	70.9	4425	10	US-09-769-097-3 Sequence 3, Appl
12	465	70.6	4317	9	US-10-044-671-1 Sequence 1, Appl
13	462	70.1	4788	10	US-09-866-866A-7 Sequence 7, Appl
14	461	70.0	3860	10	US-09-866-866A-1 Sequence 1, Appl
15	461	70.0	3860	10	US-09-866-866A-3 Sequence 3, Appl
16	461	70.0	4643	9	US-10-072-621-2 Sequence 2, Appl
17	461	70.0	8630	10	US-09-306-417-1 Sequence 1, Appl
18	461	70.0	8630	10	US-09-306-417-2 Sequence 2, Appl
19	448	68.0	3912	10	US-09-917-800A-1560 Sequence 1560, Ap
20	440	66.8	3924	10	US-09-880-107-2299 Sequence 2299, Ap
21	422	64.0	5010	10	US-09-917-800A-483 Sequence 483, App
22	369	56.0	3512	10	US-09-749-340-7 Sequence 7, Appl
23	368.5	55.9	1076	9	US-09-764-884-13 Sequence 13, Appl
24	368.5	55.9	1076	9	US-10-092-256-13 Sequence 13, Appl
25	368.5	55.9	1367	9	US-09-764-884-21 Sequence 21, Appl
26	368.5	55.9	1367	9	US-10-092-256-21 Sequence 21, Appl
27	366	55.5	4002	9	US-09-758-828-1 Sequence 1, Appl
28	366	55.5	4002	9	US-09-758-828-3 Sequence 3, Appl
29	364	55.2	1810	10	US-09-749-340-4 Sequence 4, Appl
30	364	55.2	4175	10	US-09-749-340-3 Sequence 3, Appl
31	363	55.1	3861	9	US-09-938-842A-263 Sequence 263, App
32	357	54.2	4653	9	US-10-101-388-2 Sequence 2, Appl
33	356	54.0	981	10	US-09-770-445-272 Sequence 272, App
34	350	53.1	216	10	US-09-864-761-27861 Sequence 27861, A
35	350	53.1	465	10	US-09-864-761-11237 Sequence 11237, A
36	347	52.7	939	10	US-09-974-300-6511 Sequence 6511, Ap
37	339	51.4	2681	10	US-09-749-340-8 Sequence 8, Appl
38	335	50.8	3861	9	US-09-938-842A-2237 Sequence 2237, Ap
39	333	50.5	2698	10	US-09-749-340-5 Sequence 5, Appl
40	327	49.6	53165	9	US-09-939-964-1 Sequence 1, Appl
41	326	49.5	8036	9	US-10-101-388-1 Sequence 1, Appl
42	325	49.3	2066	10	US-09-873-409-9 Sequence 9, Appl
43	319	48.4	2298	9	US-10-156-239-18 Sequence 18, Appl
44	319	48.4	2298	10	US-09-795-693-18 Sequence 18, Appl
45	319	48.4	3408	9	US-10-156-239-16 Sequence 16, Appl

ALIGNMENTS

RESULT 1
US-09-873-409-11
; Sequence 11, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
; APPLICANT: Frank, Markus
; APPLICANT: Sayegh, Mohamed
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873,409
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11
; LENGTH: 1175
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-873-409-11

Alignment Scores:
Pred. No.: 1.58e-81 Length: 1175
Score: 659.00 Matches: 131
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-873-409-3 (1-131) x US-09-873-409-11 (1-1175)

QY 1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
Db 418 ATGGTGGATGAGAAATGACATCAGAGCTTTAAATGTGGGCATTATCGAGACCATATTGGA 477
QY 21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
Db 478 GTGGTTAGTCAAGAGCCCTGTTTGTTCGGGACCCACCATCATAGTAAATATCAAGTATGGA 537
QY 41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
Db 538 CGAGATGATGTGACTGATGAAGAGATCGAGAGAGCAGCAAGGGAAGCAAAATGCGCTATGAT 597
QY 61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80
Db 598 TTTATCATGAGATTCTCTAATAATTTAAATGATTTGCTGCTGCTTAGTTTCAAAACCCCAAGATT 657
QY 81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100
Db 658 AGTGGAGGCGAGAAACAGAGATCGCAATTTGCTGCTGCTTAGTTTCAAAACCCCAAGATT 717
QY 101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120
Db 718 CTGATTTTAGATGAGGCTAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 777
QY 121 AlaLeuGluLysAspThrProArgTyrSerPhe 131
Db 778 GCACCTGGAGGAGGATACCCCAAGGATTATTCATTT 810

RESULT 2

US-09-873-409-16

; Sequence 16, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
; APPLICANT: Frank, Markus
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873, 409
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 16
; LENGTH: 1940
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-873-409-16

Alignment Scores:
Pred. No.: 3, 21e-81 Length: 1940
Score: 659.00 Matches: 131
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

US-09-873-409-3 (1-131) x US-09-873-409-16 (1-1940)

QY 1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
Db 1183 ATGGTGGATGAGAAATGACATCAGAGCTTTAAATGTGGGCATTATCGAGACCATATTGGA 1242
QY 21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
Db 1243 GTGGTTAGTCAAGAGCCCTGTTTGTTCGGGACCCACCATCATAGTAAATATCAAGTATGGA 1302
QY 41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
Db 1303 CGAGATGATGTGACTGATGAAGAGATCGAGAGAGCAGCAAGGGAAGCAAAATGCGCTATGAT 1362
QY 61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80

Db 1363 TTTATCATGAGCTTCTCTAATAATTTAAATGATTTGCTGCTGCTTAGTTTCAAAACCCCAAGATT 1422
QY 81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100
Db 1423 AGTGGAGGCGAGAAACAGAGGATCGCAATTTGCTGCTGCTTAGTTTCAAAACCCCAAGATT 1482
QY 101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120
Db 1483 CTGATTTTAGATGAGGCTAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1542
QY 121 AlaLeuGluLysAspThrProArgTyrSerPhe 131
Db 1543 GCACCTGGAGGAGGATACCCCAAGGATTATTCATTT 1575

RESULT 3

US-09-873-409-15

; Sequence 15, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
; APPLICANT: Frank, Markus
; APPLICANT: Sayegh, Mohamed
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873, 409
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 15
; LENGTH: 2021
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: No. US20020037522A1e
; LOCATION: (723)..(723)
; OTHER INFORMATION: n at position 723 represents any nucleotide (A, T, C or G)
US-09-873-409-15

Alignment Scores:

Pred. No.: 3, 4e-81 Length: 2021
Score: 659.00 Matches: 131
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

US-09-873-409-3 (1-131) x US-09-873-409-15 (1-2021)

QY 1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
Db 1264 ATGGTGGATGAGAAATGACATCAGAGCTTTAAATGTGGGCATTATCGAGACCATATTGGA 1323
QY 21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
Db 1324 GTGGTTAGTCAAGAGCCTGTTTGTTCGGGACCCACCATCATAGTAAATATCAAGTATGGA 1383
QY 41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
Db 1384 CGAGATGATGTGACTGATGAAGAGATCGAGAGAGCAGCAAGGGAAGCAAAATGCGTATGAT 1443
QY 61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80
Db 1444 TTTATCATGAGCTTCTCTAATAATTTAAATGATTTGCTGCTGCTTAGTTTCAAAACCCCAAGATT 1503
QY 81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100
Db 1504 AGTGGAGGCGAGAAACAGAGGATCGCAATTTGCTGCTGCTTAGTTTCAAAACCCCAAGATT 1563
QY 101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120
Db 1564 CTGATTTTAGATGAGGCTAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1623

QY 121 AlaLeuGluLysAspThrProArgTyrSerPhe 131
|||||
Db 1624 GCACGGAGAGGATACCCCGAGTATTCATT 1656

RESULT 4

US-09-873-409-10

; Sequence 10, Application US/09873409

; Patent No. US20020037522A1

; GENERAL INFORMATION:

; APPLICANT: Frank, Markus

; APPLICANT: Sayegh, Mohamed

; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein

; FILE REFERENCE: 81994/268611

; CURRENT APPLICATION NUMBER: US/09/873,409

; CURRENT FILING DATE: 2001-06-05

; NUMBER OF SEQ ID NOS: 19

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 10

; LENGTH: 2856

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-873-409-10

Alignment Scores:

Pred. No.:	1.84e-75	Length:	2856
Score:	619.00	Matches:	124
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	93.93%	Indels:	0
DB:	10	Gaps:	0

US-09-873-409-3 (1-131) x US-09-873-409-10 (1-2856)

QY 1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20

Db 418 ATGGTGGATGAGAAATGACATCAGAGCTTTAAATGTGCGGCATTTATCGAGACCATATTTGA 477

QY 21 ValValSerGlnGluProValLeuPheGlyThrIleSerAsnAsnIleLysTyrGly 40

Db 478 GTGGTTAGTCAGAGCTGTTTGTTCGGGACCACCATCATGTAACATATCAAGTATGA 537

QY 41 ArgAspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60

Db 538 CGAGATGATGTGACTGATGAAGAGATGGAGAGCAGCAGCAAGGCAAGCAATGCGTATGAT 597

QY 61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80

Db 598 TTTATCATGAGGTTTCTTAATAAATTTAATACATTTGGTAGGGGAAAAGGAGCTCAAAATG 657

QY 81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100

Db 658 AGTGGAGGGCAGAAACAGAGGATCGCAATTCGTGCGCTTAGTTTGAACCCCAAGATT 717

QY 101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120

Db 718 CTGATTTTAGATGAGCTACGTCTGCCCTGGATTTCAGAAAGCAAGTCAGCTGTTCAAGCT 777

QY 121 AlaLeuGluLys 124

Db 778 GCACCTGGAGAAG 789

RESULT 5

US-09-873-409-12

; Sequence 12, Application US/09873409

; Patent No. US20020037522A1

; GENERAL INFORMATION:

; APPLICANT: Frank, Markus

; APPLICANT: Sayegh, Mohamed

; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein

; FILE REFERENCE: 81994/268611

; CURRENT APPLICATION NUMBER: US/09/873,409

; CURRENT FILING DATE: 2001-06-05

; NUMBER OF SEQ ID NOS: 19

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 12

; LENGTH: 3177

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: No. US20020037522A1e

; LOCATION: (198)..(198)

; OTHER INFORMATION: n at position 198 represents any nucleotide (A, T, C or G)

US-09-873-409-12

Alignment Scores:

Pred. No.:	2.14e-75	Length:	3177
Score:	619.00	Matches:	124
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	93.93%	Indels:	0
DB:	10	Gaps:	0

US-09-873-409-3 (1-131) x US-09-873-409-12 (1-3177)

QY 1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20

Db 739 ATGGTGGATGAGAAATGACATCAGAGCTTTAAATGTGCGGCATTTATCGAGACCATATTTGA 798

QY 21 ValValSerGlnGluProValLeuPheGlyThrIleSerAsnAsnIleLysTyrGly 40

Db 799 GTGGTTAGTCAGAGCTGTTTGTTCGGGACCACCATCATGTAACATATCAAGTATGA 858

QY 41 ArgAspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60

Db 859 CCAGATGATGTGACTGATCAAGAGATGGAGAGCAGCAGCAAGGCAAGCAATGCGTATGAT 918

QY 61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80

Db 919 TTTATCATGAGGTTTCTTAATAAATTTAATACATTTGGTAGGGGAAAAGGAGCTCAAAATG 978

QY 81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100

Db 979 AGTGGAGGGCAGAAACAGAGGATCGCAATTCGTGCGCTTAGTTTGAACCCCAAGATT 1038

QY 101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120

Db 1039 CTGATTTTAGATGAGCTACGTCTGCCCTGGATTTCAGAAAGCAAGTCAGCTGTTCAAGCT 1098

QY 121 AlaLeuGluLys 124

Db 1099 GCACCTGGAGAAG 1110

RESULT 6

US-09-873-409-14

; Sequence 14, Application US/09873409

; Patent No. US20020037522A1

; GENERAL INFORMATION:

; APPLICANT: Frank, Markus

; APPLICANT: Sayegh, Mohamed

; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein

; FILE REFERENCE: 81994/268611

; CURRENT APPLICATION NUMBER: US/09/873,409

; CURRENT FILING DATE: 2001-06-05

; NUMBER OF SEQ ID NOS: 19

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 14

; LENGTH: 3621

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-873-409-14

Alignment Scores:

Pred. No.:	2.57e-75	Length:	3621
------------	----------	---------	------

Score: 619.00 Matches: 124
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 93.93% Indels: 0
DB: 10 Gaps: 0

US-09-873-409-3 (1-131) x US-09-873-409-14 (1-3621)

QY 1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
DB 1183 ATGGTGGATGAGATGACATCAGAGCTTTTAAATGTGGCGCATTCGAGACCATATTGGA 1242
QY 21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
DB 1243 GTGGTTAGTCAAGAGCCTGTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGA 1302
QY 41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
DB 1303 CGAGATGATGTGCTGATGAAGAGATGGAGAGCAGCAAGGGAAGCAATGCGTATGAT 1362
QY 61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80
DB 1363 TTTATCATGGAGTTTCTTAATAATTTAATACATTTGTAGGGGAAAAGAGCTCAAAATG 1422
QY 81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100
DB 1423 AGTGGAGGCGAGAAACAGAGATCGCAATTTGCTCGTCCCTAGTTCGAAACCCCAAGATT 1482
QY 101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120
DB 1483 CTGATTTTAGATGAGGCTAGCTGTGCTCCCTGGATTTCAGAAAGCAAGTCAGCTGTTCAGCT 1542
QY 121 AlaLeuGluLys 124
DB 1543 GCACCTGGAGAAG 1554

RESULT 7

US-09-873-409-13
; Sequence 13, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
; APPLICANT: Frank, Markus
; APPLICANT: Sayegh, Mohamed
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
; TITLE OF INVENTION: Homologue on Chromosome 7p15-21 and Uses Thereof
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873,409
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 13
; LENGTH: 3702
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: No. US20020037522A1e
; LOCATION: (723) .. (723)
; OTHER INFORMATION: n at position 723 represents any nucleotide (A, T, C or G)
US-09-873-409-13

Alignment Scores:

Pred. No.: 2,65e-75 Length: 3702
Score: 619.00 Matches: 124
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 93.93% Indels: 0
DB: 10 Gaps: 0

US-09-873-409-3 (1-131) x US-09-873-409-13 (1-3702)

QY 1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
DB 1264 ATGGTGGATGAGATGACATCAGAGCTTTTAAATGTGGCGCATTCGAGACCATATTGGA 1323

QY 21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
DB 1324 GTGGTTAGTCAAGAGCCTGTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGA 1383
QY 41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
DB 1384 CGAGATGATGTGCTGATGAAGAGATGGAGAGCAGCAAGGGAAGCAATGCGTATGAT 1443
QY 61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80
DB 1444 TTTATCATGGAGTTTCTTAATAATTTAATACATTTGTAGGGGAAAAGAGCTCAAAATG 1503
QY 81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100
DB 1504 AGTGGAGGCGAGAAACAGAGATCGCAATTTGCTCGTCCCTAGTTCGAAACCCCAAGATT 1563
QY 101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120
DB 1564 CTGATTTTAGATGAGGCTAGCTGTGCTCCCTGGATTTCAGAAAGCAAGTCAGCTGTTCAGCT 1623
QY 121 AlaLeuGluLys 124
DB 1624 GCACCTGGAGAAG 1635

RESULT 8

US-09-866-866A-5
; Sequence 5, Application US/09866866A
; Patent No. US20020102244A1
; GENERAL INFORMATION:
; APPLICANT: Sorrentino, Brian
; APPLICANT: Schuetz, John
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
; FILE REFERENCE: 1340-1-021CIP2
; CURRENT APPLICATION NUMBER: US/09/866,866A
; CURRENT FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 09/584,586
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: PCT/US99/11825
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: 60/086,988
; PRIOR FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 4189
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-866-866A-5

Alignment Scores:

Pred. No.: 2.2e-54 Length: 4189
Score: 468.00 Matches: 90
Percent Similarity: 91.06% Conservative: 22
Best Local Similarity: 73.17% Mismatches: 11
Query Match: 71.02% Indels: 0
DB: 10 Gaps: 0

US-09-873-409-3 (1-131) x US-09-866-866A-5 (1-4189)

QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
DB 1354 ATCGACGGGCAAGACATCAGAACCATCAATGTGAGGTATCTGAGGGAGATCAATTTGGTGTG 1413
QY 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
DB 1414 GTGAGTCAGGAACCTGTGCTTTGCCACCATCGATCGCCGAGAACATTCGCTATGCGCGA 1473
QY 42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
DB 1474 GAGATGTCCCATGATGAGATTGAGAAGCTGTCAAGGAGCCCAATGCTATGACTTC 1533
QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81


```
QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyLysGlyAlaGlnMetSer 81
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1864 ATCATGAACCTGCCACCAATTTGACACCTTGTTGGTGAGAGGGCGCAGCTGAGT 1923

QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1924 GGGGACAGAAACAGAGGATCGCATTTGCCGGGCGCTGTCGCCCAACCAAGATCCTT 1983

QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1984 TTGTGGATGAGCCACGTCAGCTTGGACACAGAAAGCAAGCGGTTCAGGCGCGCT 2043

QY 122 LeuGluLys 124
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2044 CTGGATAAG 2052

RESULT 11
US-09-769-097-3
; Sequence 3, Application US/09769097
; Patent No. US20020055128A1
; GENERAL INFORMATION:
; APPLICANT: Kimberly Anne Brun
; APPLICANT: Richard James Chenery
; APPLICANT: Harma Ellens
; APPLICANT: John Anthony Feild
; APPLICANT: Lin Yue
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES
; FILE REFERENCE: GP-50009-C2
; CURRENT APPLICATION NUMBER: US/09/769,097
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 09/208,809
; PRIOR FILING DATE: 1998-12-09
; PRIOR APPLICATION NUMBER: 09/156,800
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: US99/20770
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 4425
; TYPE: DNA
; ORGANISM: RATTUS RATTUS
US-09-769-097-3

Alignment Scores:
Pred. No.: 3,27e-54 Length: 4425
Score: 467.00 Matches: 90
Percent Similarity: 91.06% Conservative: 22
Best Local Similarity: 73.17% Mismatches: 11
Query Match: 70.86% Indels: 0
DB: 10 Gaps: 0

US-09-873-409-3 (1-131) x US-09-769-097-3 (1-4425)
QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1684 ATTCACCGACAGGACATCAGACCATCAATGTAGGTATCTCGGGAAATCATTTGGGGTG 1743

QY 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1744 GTGAGTCAGGAACCCGTCGTTTGGCCACACCAATTCGCCGAAACATTCGCTATGCGCGA 1803

QY 42 AspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1804 GAAACCTGACCATGATGATAGATAAGAAAGCTGTCAAGGAAGCAATGCTATGATTTC 1863

QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyLysGlyAlaGlnMetSer 81
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1864 ATCATGAACCTGCCACCAATTTGACACCTTGTTGGTGAGAGGGCGCAGCTGAGT 1923

QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1923
```

```
Db 1924 GGGGACAGAAACAGAGGATCCCATTCGCCGGGCGCTGTCGCCCAACCAAGATCCTT 1983

QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1984 TTGTGGATGAGCCACGTCAGCTTGGACACAGAAAGCAAGCGGTTCAGGCGCGCT 2043

QY 122 LeuGluLys 124
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2044 CTGGATAAG 2052

RESULT 12
US-10-044-671-1
; Sequence 1, Application US/10044671
; Patent No. US20020177147A1
; GENERAL INFORMATION:
; APPLICANT: Washington State University Research Foundation
; APPLICANT: Mealey, Katrina
; APPLICANT: Bentjen, Steven
; TITLE OF INVENTION: MDRI VARIANTS AND METHODS FOR THEIR USE
; FILE REFERENCE: 4630-61733
; CURRENT APPLICATION NUMBER: US/10/044,671
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/261,578
; PRIOR FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: US 60/314,829
; PRIOR FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 4317
; TYPE: DNA
; ORGANISM: Canis familiaris
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (70)..(3912)
; OTHER INFORMATION:
US-10-044-671-1

Alignment Scores:
Pred. No.: 5,96e-54 Length: 4317
Score: 465.00 Matches: 89
Percent Similarity: 89.43% Conservative: 21
Best Local Similarity: 72.36% Mismatches: 13
Query Match: 70.56% Indels: 0
DB: 9 Gaps: 0

US-09-873-409-3 (1-131) x US-10-044-671-1 (1-4317)
QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1429 ATTGATGGACAGACATTAGGACCATTAATGTAAGGCATCTTCGGGAATTAAGTGGTG 1488

QY 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1489 GTGAGTCAGGAGCCTGTTTGGCCACACCATAGCTGAAACATTCGCTATGCGCGC 1548

QY 42 AspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1549 GAAATGTCCACCATGATGAGATTGAAAGCTGTTAAGGAAGCAATGCTATGATTTT 1608

QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyLysGlyAlaGlnMetSer 81
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1609 ATCATGAACCTACCTAATAATTTGACACTCTGGTTGGAGAGAGGGCGCGCTGAGT 1668

QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1669 GGTGGACAGAAACAGAGAATCGCCATTCGTCGGGCGCTGTTGTCGCAACCAAGATTC 1728

QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1729 CTGCTGGATGAGGCAACGTCAGCTCTGGACACTGGAAGTGAAGCAGTGGTTTCAAGTGGCC 1788
```


QY 122 LeuGlulys 124
|||:|||||
Db 1789 CTGGATAAG 1797

RESULT 13

US-09-866-866A-7

; Sequence 7, Application US/09866866A

; Patent No. US20020102244A1

; GENERAL INFORMATION:

; APPLICANT: Sorrentino, Brian

; APPLICANT: Schuetz, John

; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells

; FILE REFERENCE: 1340-1-021CIP2

; CURRENT APPLICATION NUMBER: US/09/866,866A

; CURRENT FILING DATE: 2001-08-30

; PRIOR FILING DATE: 2001-08-30

; PRIOR APPLICATION NUMBER: 09/584,586

; PRIOR FILING DATE: 2000-05-31

; PRIOR APPLICATION NUMBER: PCT/US99/11825

; PRIOR FILING DATE: 1999-05-27

; PRIOR APPLICATION NUMBER: 60/086,988

; NUMBER OF SEQ ID NOS: 27

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 7

; LENGTH: 4788

; TYPE: DNA

; ORGANISM: Mus musculus

; US-09-866-866A-7

Alignment Scores:

Pred. No.:	1.79e-53	Length:	4788
Score:	462.00	Matches:	89
Percent Similarity:	90.24%	Conservative:	22
Best Local Similarity:	72.36%	Mismatches:	12
Query Match:	70.11%	Indels:	0
DB:	10	Gaps:	0

US-09-873-409-3 (1-131) x US-09-866-866A-7 (1-4788)

QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
|||:|||||

Db 1345 ATCCAGCGGACAGACATCAAGAACATCATGTGAGTAATCTGAGGAGATCATTTGGTGTG 1404
|||:|||||

QY 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
|||:|||||

Db 1405 GTGAGTCAGGAACCTGTGCTGTTTGCACACGATCGCCGAGAACATTCGCTATGCCCGA 1464
|||:|||||

QY 42 AspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
|||:|||||

Db 1465 GAAGATGTCCACATGGATGAGATGTGAGAAAGCTGTCAAGGAGCCCAATGCCTATGACTTC 1524
|||:|||||

QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
|||:|||||

Db 1525 ATCATGAACTGCCCAACCAATTGACACCTGTTGGTGAGAGAGGGCGACGTGAGT 1584
|||:|||||

QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
|||:|||||

Db 1585 GGGGACAGAAACAGAGATCGCATTCGCCGGCCCTGCTCCGCAATCCCAAGATCCTT 1644
|||:|||||

QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
|||:|||||

Db 1645 TTGTTGGAGAGGCCACCTCAGCCCTGGATACAGAAAGTGAAGCTGTGTTTCAGGCCGCA 1704
|||:|||||

QY 122 LeuGlulys 124
|||:|||||

Db 1705 CTGGATAAG 1713

RESULT 14

US-09-866-866A-1

; Sequence 1, Application US/09866866A

; Patent No. US20020102244A1

; GENERAL INFORMATION:

; APPLICANT: Sorrentino, Brian

; APPLICANT: Schuetz, John

; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells

; FILE REFERENCE: 1340-1-021CIP2

; CURRENT APPLICATION NUMBER: US/09/866,866A

; CURRENT FILING DATE: 2001-08-30

; PRIOR FILING DATE: 2001-08-30

; PRIOR APPLICATION NUMBER: 09/584,586

; PRIOR FILING DATE: 2000-05-31

; PRIOR APPLICATION NUMBER: PCT/US99/11825

; PRIOR FILING DATE: 1999-05-27

; PRIOR APPLICATION NUMBER: 60/086,988

; APPLICANT: Schuetz, John
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
; FILE REFERENCE: 1340-1-021CIP2
; CURRENT APPLICATION NUMBER: US/09/866,866A
; CURRENT FILING DATE: 2001-08-30
; PRIOR FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 09/584,586
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: PCT/US99/11825
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: 60/086,988
; PRIOR FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 3860
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-866-866A-1

Alignment Scores:

Pred. No.:	1.82e-53	Length:	3860
Score:	461.00	Matches:	90
Percent Similarity:	89.43%	Conservative:	20
Best Local Similarity:	73.17%	Mismatches:	13
Query Match:	69.95%	Indels:	0
DB:	10	Gaps:	0

US-09-873-409-3 (1-131) x US-09-866-866A-1 (1-3860)

QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
|||:|||||

Db 1357 GTTGATGGACAGATATTAGGACCATTAATGTAAAGTTTCTACGGGAAATCATTTGGTGTG 1416
|||:|||||

QY 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
|||:|||||

Db 1417 GTGAGTCAGGAACCTGTATTGTTTGCACACGATAGCTGAAACACATTCGCTATGCCCGT 1476
|||:|||||

QY 42 AspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
|||:|||||

Db 1477 GAAATGTCCACATGGATGAGATGTGAGAAAGCTGTCAAGGAGCCCAATGCCTATGACTTT 1536
|||:|||||

QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
|||:|||||

Db 1537 ATCATGAACTGCCCAACCAATTGACACCTGTTGGTGAGAGAGGGCGCCAGTTGAGT 1596
|||:|||||

QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
|||:|||||

Db 1597 GGTGGCAGACAGAGGATCCCATTCGACCTGCTGTTGGACACCCCAAGATCCTC 1656
|||:|||||

QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
|||:|||||

Db 1657 CTGCTGATGAGCCACCTGAGCCCTTGGACACAGAAAGCGACAGTGTTCAGGTGGCT 1716
|||:|||||

QY 122 LeuGlulys 124
|||:|||||

Db 1717 CTGGATAAG 1725

RESULT 15

US-09-866-866A-3

; Sequence 3, Application US/09866866A

; Patent No. US20020102244A1

; GENERAL INFORMATION:

; APPLICANT: Sorrentino, Brian

; APPLICANT: Schuetz, John

; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells

; FILE REFERENCE: 1340-1-021CIP2

; CURRENT APPLICATION NUMBER: US/09/866,866A

; CURRENT FILING DATE: 2001-08-30

; PRIOR FILING DATE: 2001-08-30

; PRIOR APPLICATION NUMBER: 09/584,586

; PRIOR FILING DATE: 2000-05-31

; PRIOR APPLICATION NUMBER: PCT/US99/11825

; PRIOR FILING DATE: 1999-05-27

; PRIOR APPLICATION NUMBER: 60/086,988

; PRIOR FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 3860
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-866-866A-3

Alignment Scores:
Pred. No.: 1.82e-53 Length: 3860
Score: 461.00 Matches: 90
Percent Similarity: 89.43% Conservative: 20
Best Local Similarity: 73.17% Mismatches: 13
Query Match: 69.95% Indels: 0
DB: 10 Gaps: 0

US-09-873-409-3 (1-131) x US-09-866-866A-3 (1-3860)

```
QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
Db 1357 GTTGATGCAGCAGATATTAGACCATAAATGTAAGTTTCTACGGGAATCATTTGGTG 1416

QY 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
Db 1417 GTGAGTCAGGAACCTGTATTGTTTGCACCACCATAGCTGAAAAACATTCGCTATGGCGGT 1476

QY 42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
Db 1477 GAAATGTCACCATGATGAGATTGAGAAAGCTGTCAAGGAAGCCATGCTATGACTTT 1536

QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
Db 1537 ATCATGAACCTGCTCATAAATTGTACACCTGGTTGGAGAGAGGGGCCAGTTGAGT 1596

QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
Db 1597 GGTGGGCAGAAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTGCAACCCCAAGATCCTC 1656

QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
Db 1657 CTGCTGGATGAGGCCACGTGAGCCTTGGACACAGAAAGCGAGTGTTTCAGGTGGCT 1716

QY 122 LeuGluLys 124
Db 1717 CTGGATAAG 1725
```

Search completed: March 31, 2003, 16:00:29
Job time : 58.683 secs

Result No.	Score	Query %		Length	DB	ID	Description
		Match					
1	472	71.6	4233	3	US-09-120-513-1		Sequence 1, Appli
2	472	71.6	4233	4	US-09-450-105-1		Sequence 1, Appli
3	461	70.0	4284	2	US-08-784-649A-1		Sequence 1, Appli
4	461	70.0	4264	2	US-08-784-649A-5		Sequence 5, Appli
5	461	70.0	4646	1	US-08-181-471-2		Sequence 2, Appli
6	461	70.0	4669	2	US-08-752-447-1		Sequence 1, Appli
7	461	70.0	4659	4	US-09-316-167-1		Sequence 1, Appli
8	461	70.0	4659	6	5206352-3		Patent No. 5206352
9	461	70.0	6505	2	US-08-793-610-5		Sequence 5, Appli
10	461	70.0	9318	2	US-08-793-610-6		Sequence 6, Appli
11	449	68.1	4669	2	US-08-583-276-18		Sequence 18, Appli
12	372	56.4	2736	1	US-08-461-823-1		Sequence 1, Appli

/ TELEFAX: 610-270-5090
/ TELEX:
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 4233 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: Genomic DNA
US-09-120-513-1

Alignment Scores:
Pred. No.: 3.59e-55 Length: 4233
Score: 472.00 Matches: 91
Percent Similarity: 91.06% Conservatives: 21
Best Local Similarity: 73.98% Mismatches: 11
Query Match: 71.62% Indels: 0
DB: 3 Gaps: 0

US-09-873-409-3 (1-131) x US-09-120-513-1 (1-4233)

QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
DB 1376 ATCGACGACGACGACATCAGGACCAATGTGAGGTATCTGCGGGAATCATTTGGGGTG 1435
QY 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
DB 1436 GTGAGTCAGAAACCCGCTGCTTTGCCACACAGATTCGCCGAAACATTCGCTATGCGCGA 1495
QY 42 AspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
DB 1496 GAAACGTCACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 61
QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGlyAlaGlnMetSer 81
DB 1556 ATCATGAAACTGCCCCACAAATTTAAACACCTGTTGGTGAGAGGGCGGCGAGCTGAGT 1615
QY 82 GlyGlyGlnLysGlnArgAlaAlaAlaArgAlaLeuValArgAsnProLysIleLeu 101
DB 1616 GGGGACAGAAACAGAGATCGCATTTGCCGGCCCTGTCGCCAACCCCAAGATCCTT 1675
QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
DB 1676 TTGTTGATGAGGCGACGTCAGCCTTGACACAGAAAGCGCGGTTCAGGCGCGCT 1735
QY 122 LeuGluLys 124
DB 1736 CTGGATAAG 1744

RESULT 2

US-09-450-105-1
; Sequence 1, Application US/09450105
; Patent No. 6169166
; GENERAL INFORMATION:
; APPLICANT: Kimberly Anne Brun
; APPLICANT: Richard James Chenery
; APPLICANT: Hama Ellens
; APPLICANT: John Anthony Feld
; APPLICANT: Lin Yue
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES
; FILE REFERENCE: GP-50008-D1
; CURRENT APPLICATION NUMBER: US/09/450,105
; EARLIER FILING DATE: 1999-11-29
; EARLIER APPLICATION NUMBER: 09/120,513
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 4233
; TYPE: DNA
; ORGANISM: HOMO SAPIENS
US-09-450-105-1

Alignment Scores:

Pred. No.: 3.59e-55 Length: 4233
Score: 472.00 Matches: 91
Percent Similarity: 91.06% Conservatives: 21
Best Local Similarity: 73.98% Mismatches: 11
Query Match: 71.62% Indels: 0
DB: 4 Gaps: 0

US-09-873-409-3 (1-131) x US-09-450-105-1 (1-4233)

QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
DB 1376 ATCGACGACGACGACATCAGGACCAATGTGAGGTATCTGCGGGAATCATTTGGGGTG 1435
QY 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
DB 1436 GTGAGTCAGAAACCCGCTGCTTTGCCACACAGATTCGCCGAAACATTCGCTATGCGCGA 1495
QY 42 AspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
DB 1496 GAAACGTCACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 61
QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGlyAlaGlnMetSer 81
DB 1556 ATCATGAAACTGCCCCACAAATTTAAACACCTGTTGGTGAGAGGGCGGCGAGCTGAGT 1615
QY 82 GlyGlyGlnLysGlnArgAlaAlaAlaArgAlaLeuValArgAsnProLysIleLeu 101
DB 1616 GGGGACAGAAACAGAGATCGCATTTGCCGGCCCTGTCGCCAACCCCAAGATCCTT 1675
QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
DB 1676 TTGTTGATGAGGCGACGTCAGCCTTGACACAGAAAGCGCGGTTCAGGCGCGCT 1735
QY 122 LeuGluLys 124
DB 1736 CTGGATAAG 1744

RESULT 3

US-08-784-649A-1
; Sequence 1, Application US/08784649A
; Patent No. 5830697
; GENERAL INFORMATION:
; APPLICANT: Sikic, Branimir I
; APPLICANT: Chen, Gang
; TITLE OF INVENTION: P-GLYCOPROTEIN MUTANT RESISTANT TO
; NUMBER OF INVENTIONS: 5
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 2200 Sand Hill Road
; CITY: Menlo Park
; STATE: CA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/784,649A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: Reg No. 5830697 36,677
; REFERENCE/DOCKET NUMBER: 06037/007001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-322-5070
; TELEFAX: 415-854-0875
; INFORMATION FOR SEQ ID NO: 1:


```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/181,471
; FILING DATE: 13-JAN-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/041,553
; FILING DATE: 02-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: ANT0029P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-792-3680
; TELEFAX: 619-792-8477
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4646 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 425..4267
; US-08-181-471-2

Alignment Scores:
Pred. No.: 1,37e-53 Length: 4646
Score: 461.00 Matches: 90
Percent Similarity: 89.43% Conservative: 20
Best Local Similarity: 73.17% Mismatches: 13
Query Match: 69.95% Indels: 0
DB: 1 Gaps: 0

US-09-873-409-3 (1-131) x US-08-181-471-2 (1-4646)
QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
Db 1781 GTTGATGCAGCAGATATTAGGACCAATAATGTAAGGTTCTACGGGAATCAATGGTG 1840
QY 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
Db 1841 GTGAGTCAGGAACCTGTATTGTTTGGCCACCACGATAGCTGAAACATTCGCTATGCCGT 1900
QY 42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
Db 1901 GAAATGTCCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCAATGCCTATGACTTT 1960
QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
Db 1961 ATCATGAACTGCCTCATAAATTTGACACCTGTTGGAGAGAGAGGGCCAGTTGAGT 2020
QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
Db 2021 GGTGGGAGAGAGAGAGATCGCCATTCGACGTCCCTGTTGTCACACCCCAAGATCCCT 2080
QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
Db 2081 CTGCTGGATGAGGCCACGTGACGCTTGACACAGAAAGCGAGAGGTTTCAGGTGCT 2140
QY 122 LeuGluLys 124
Db 2141 CTGGATAAG 2149

RESULT 6
US-08-752-447-1
; Sequence 1, Application US/08752447
; Patent No. 5994088
; GENERAL INFORMATION:
; APPLICANT: Mechetner, Eugene
; APPLICANT: Roninson, Igor B
; TITLE OF INVENTION: Methods and Reagents for Preparing and
; TITLE OF INVENTION: Using Immunological Agents Specific for P-glycoprotein
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff Ltd.
; STREET: 300 South Wacker Drive, Seventh Floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/752,447
; FILING DATE: 15-NOV-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5994088nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 95,1121
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-9808
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4669 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: 5' UTR
; LOCATION: 1..424
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 425..4264
; FEATURE:
; NAME/KEY: 3' UTR
; LOCATION: 4265..4669
; US-08-752-447-1

Alignment Scores:
Pred. No.: 1,38e-53 Length: 4669
Score: 461.00 Matches: 90
Percent Similarity: 89.43% Conservative: 20
Best Local Similarity: 73.17% Mismatches: 13
Query Match: 69.95% Indels: 0
DB: 2 Gaps: 0

US-09-873-409-3 (1-131) x US-08-752-447-1 (1-4669)
QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
Db 1781 GTTGATGCAGCAGATATTAGGACCAATAATGTAAGGTTCTACGGGAATCAATGGTG 1840
QY 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
Db 1841 GTGAGTCAGGAACCTGTATTGTTTGGCCACCACGATAGCTGAAACATTCGCTATGCCGT 1900
QY 42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
Db 1901 GAAATGTCCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCAATGCCTATGACTTT 1960
QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
Db 1961 ATCATGAACTGCCTCATAAATTTGACACCTGTTGGAGAGAGAGGGCCAGTTGAGT 2020
QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
```


Db 2021 GGTGGCAGACAGAGGATCGCATTCGACGTCCTGTTGCGAACCCCAAGATCCTC 2080
Qy 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerIleValGlnAlaA 121
Db 2081 CTGCTGATGAGCCACGTCAGCTTGGACACAGAAAGCGACGCTGGTTCAAGTGGCT 2140
Qy 122 LeuGluLys 124
Db 2141 CTGGATAAG 2149
RESULT 7
US-09-316-167-1
; Sequence 1, Application US/09316167
; Patent No. 6365357
; GENERAL INFORMATION:
; APPLICANT: Mechtner, Eugene
; APPLICANT: Roninson, Igor B
; TITLE OF INVENTION: Methods and Reagents for Preparing and
; USING IMMUNOLOGICAL AGENTS SPECIFIC FOR P-GLYCOPROTEIN
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boenken Hulbert & Berghoff Ltd.
; STREET: 300 South Wacker Drive, Seventh Floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/316,167
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/752,447
; FILING DATE: 15-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6365357man, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 95,1121
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-9808
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4669 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: 1..424
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 425..4264
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: 4265..4669
US-09-316-167-1
Alignment Scores:
Pred. No.: 1.38e-53 Length: 4669
Score: 461.00 Matches: 90
Percent Similarity: 89.43% Conservative: 20
Best Local Similarity: 73.17% Mismatches: 13
Query Match: 69.95% Indels: 0
DB: 4 Gaps: 0

US-09-873-409-3 (1-131) x US-09-316-167-1 (1-4669)
Qy 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
Db 1781 GTTGATGGACAGCATATTAGGACCATAAATGTAAGTTTCTACGGGAAATCATTTGGTGTG 1840
Qy 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
Db 1841 GTGAGTCAGGAACCTGTATTGTTGCCACCATGATAGCTGAAACAAATTCGCTATGGCCGT 1900
Qy 42 AspAspValThrAspGluMetGluArgAlaAaArgGluAlaAsnAlaTyrAspPhe 61
Db 1901 GAAATGTCCATGATGAGATTGAGAAAGCTGTCAAGGAAGCAATCCCTATGACTTT 1960
Qy 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
Db 1961 ATCATGAAACTGCCTCATAAATTTGACACCTGTTGGAGAGAGAGGGGCCCAAGTTGAGT 2020
Qy 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
Db 2021 GGTGGCAGAACGAGGATGCCATTGCACCTGTTGGAGAGAGAGGGGCCCAAGATCCTC 2080
Qy 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerIleValGlnAlaAa 121
Db 2081 CTGCTGATGAGCCACGTCAGCTTGGACACAGAAAGCGACGCTGGTTCAAGTGGCT 2140
Qy 122 LeuGluLys 124
Db 2141 CTGGATAAG 2149
RESULT 8
5206352-3
; Patent No. 5206352
; APPLICANT: Roninson, Igor B.; Pastan Ira H.; Gottesman,
; Michael M.
; TITLE OF INVENTION: COMPOSITIONS FOR CLONES CONTAINING DNA
; SEQUENCES ASSOCIATED WITH MULTIDRUG RESISTANCE IN HUMAN CELLS
; NUMBER OF SEQUENCES: 4
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/622,836
; FILING DATE: 24-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 892,575
; FILING DATE: 01-AUG-1986
; APPLICATION NUMBER: 845,610
; FILING DATE: 28-MAR-1986
; SEQ ID NO:3
; LENGTH: 4669
5206352-3
Alignment Scores:
Pred. No.: 1.38e-53 Length: 4669
Score: 461.00 Matches: 90
Percent Similarity: 89.43% Conservative: 20
Best Local Similarity: 73.17% Mismatches: 13
Query Match: 69.95% Indels: 0
DB: 6 Gaps: 0
US-09-873-409-3 (1-131) x 5206352-3 (1-4669)
Qy 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
Db 1781 GTTGATGGACAGCATATTAGGACCATAAATGTAAGTTTCTACGGGAAATCATTTGGTGTG 1840
Qy 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
Db 1841 GTGAGTCAGGAACCTGTATTGTTGCCACCATGATAGCTGAAACAAATTCGCTATGGCCGT 1900
Qy 42 AspAspValThrAspGluMetGluArgAlaAaArgGluAlaAsnAlaTyrAspPhe 61
Db 1901 GAAATGTCCATGATGAGATTGAGAAAGCTGTCAAGGAAGCAATCCCTATGACTTT 1960
Qy 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81


```
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9318 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA
US-08-793-610-6

Alignment Scores:
Pred. No.: 3 7e-53 Length: 9318
Score: 461.00 Matches: 90
Percent Similarity: 89.43% Conservative: 20
Best Local Similarity: 73.17% Mismatches: 13
Query Match: 69.95% Indels: 0
DB: 2 Gaps: 0

US-09-873-409-3 (1-131) x US-08-793-610-6 (1-9318)
QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
Db 3132 GTTGATGGACAGCATATTAGGACCAATAATGTAAGGTTCTACGGGAATCATTTGGTGTG 3191
QY 22 ValSerGlnGluProValLeuPheGlyThrIleSerAsnAsnIleLysTyrGlyArg 41
Db 3192 GTGAGTCAGGAACCTGTATTGTTTGCACCAACGATAGCTGAAAACATTCGCTATGCGCGT 3251
QY 42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
Db 3252 GAAATGTCCACCATGATGATGATGAGAAAGCTGTCAAGGAAGCAATGCTATGACTTT 3311
QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
Db 3312 ATCATGAACTGCTCATATAATTGACACCTGTGTTGAGAGAGAGGGGCCAGTTGAGT 3371
QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
Db 3372 GGTGGGAGAGAGAGAGATCGCCATTCACGTCGCTGTTGCAACCAAGATGCTC 3431
QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaA 121
Db 3432 CTGCTGGATGAGGCCACGTGAGCCTTTGGACACAGAAAGCGAGCGTTCAGGTGCGT 3491
QY 122 LeuGluLys 124
Db 3492 CTGGATAAG 3500

RESULT 11
US-08-583-276-18
; Sequence 18, Application US/08583276
; Patent No. 5837536
; GENERAL INFORMATION:
; APPLICANT: McDonagh, Kevin T.
; APPLICANT: Nienhuis, Arthur
; APPLICANT: Tolstoshev, Paul
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HUMAN
; TITLE OF INVENTION: MULTIDRUG RESISTANCE GENES AND IMPROVED
; TITLE OF INVENTION: SELECTION OF CELLS TRANSFECTED WITH SUCH GENES
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan,
; ADDRESSEE: Cecchi & Stewart
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: DW4.V2
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/583,276
; FILING DATE: 05-JAN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/332,444
; FILING DATE: 31-OCT-1994
; APPLICATION NUMBER: 07/887,712
; FILING DATE: 22-MAY-1992
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4669 bases
; TYPE: nucleic acid
; STRANDEDNESS: singular
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: Genomic DNA
US-08-583-276-18

Alignment Scores:
Pred. No.: 6 31e-52 Length: 4669
Score: 449.00 Matches: 88
Percent Similarity: 87.80% Conservative: 20
Best Local Similarity: 71.54% Mismatches: 15
Query Match: 68.13% Indels: 0
DB: 2 Gaps: 0

US-09-873-409-3 (1-131) x US-08-583-276-18 (1-4669)
QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
Db 1781 GTTCATGGACAGCATATTAGGACCAATAATGTAAGGTTCTACGGGAATCATTTGGTGTG 1840
QY 22 ValSerGlnGluProValLeuPheGlyThrIleSerAsnAsnIleLysTyrGlyArg 41
Db 1841 GTGAGTCAGGAACCTGTATTGTTGCCACCAACGATAGCTGAAAACATTCGCTATGCGCGT 1900
QY 42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
Db 1901 GAAATGTCCACCATGATGATGAGAAAGCTGTCAAGGAAGCAATGCTATGACTTT 1960
QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
Db 1961 ATCATGAACTGCTCATATAATTGACACCTGTGTTGAGAGAGAGGGGCCAGTTGAGT 2020
QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
Db 2021 GGTGGGAGAGAGAGATCGCCATTCACGTCGCTGTTGCAACCAAGATGCTC 2080
QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaA 121
Db 2081 CTGCTGGATGAGGCCACGTGAGCCTTTGGACACAGAAAGCGAGCGTTCAGGTGCGT 2140
QY 122 LeuGluLys 124
Db 2141 CTGGATAAG 2149

RESULT 12
US-08-461-823-1
; Sequence 1, Application US/08461823
; Patent No. 5593840
; GENERAL INFORMATION:
; APPLICANT: Bhatnagar, Satish K.
; APPLICANT: George Jr., Albert L.
; APPLICANT: Nazarenko, Irina
; TITLE OF INVENTION: AMPLIFICATION OF NUCLEIC ACID SEQUENCES
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OncorPharm, Inc.
; STREET: 200 Perry Parkway
; CITY: Gaithersburg
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20877
```



```
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/461,823
/ FILING DATE: 05-JUN-1995
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/168,621
/ FILING DATE: 16-DEC-1993
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/010,433
/ FILING DATE: 27-JAN-1993
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Karta, Glenn E.
/ REGISTRATION NUMBER: 30,649
/ REFERENCE/DOCKET NUMBER: PA-0012 CIP 2
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 301 527-2058
/ TELEFAX: 301 208-6997
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 2726 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ US-08-461-823-1

Alignment Scores:
Pred. No.: 1.33e-41 Length: 2726
Score: 372.00 Matches: 75
Percent Similarity: 77.78% Conservative: 23
Best Local Similarity: 59.52% Mismatches: 26
Query Match: 56.45% Indels: 2
DB: 1 Gaps: 1

US-09-873-409-3 (1-131) x US-08-461-823-1 (1-2726)
QY 1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
Db 1787 CTGCTTCATGGCAAGAAATAAAGCGACTGAATGTTGAGTGGCTCCGAGCACACCTGGGC 1846
QY 21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
Db 1847 ATCGTGCCAGGAGCCCATCTGTTGACTGCAGCATTCGTGAGAACATTCCTATGGA 1906
QY 41 ArgAspAsp-----ValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAla 58
Db 1907 GACAAAGCCCGGGTGTGTACAGAGAGAGATCGTGAGGCGCAGCAAGAGGAGGCCAACA 1966
QY 59 TyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAla 78
Db 1967 CATGCCCTTCATCGAGTCAGCTCAGCTTAATAATATATAGCACTAAGTAGAGACAAAGGAACT 2026
QY 79 GlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnPro 98
Db 2027 CAGCTCTCTGTGGCCAGAAACACGACCATTCGCCATAGCTCGTGCCTTGTGTAGACGCT 2086
QY 99 LysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaVal 118
Db 2087 CATATTTTGTCTTTGGATGAAGCCACGTCAGCTCTGGATACAGAAAGTGAAGAGTTGTC 2146
QY 119 GlnAlaAlaLeuGluLys 124
Db 2147 CAGAAGCCCTTGGACAAA 2164
RESULT 13
US-08-996-545-1

/
/ Sequence 1, Application US/08996545
/ Patent No. 5928898
/ GENERAL INFORMATION:
/ APPLICANT: Skatrud, Paul L.
/ APPLICANT: de Waard, Maarten A.
/ APPLICANT: Peery, Robert B.
/ APPLICANT: Andrade, Alan C.
/ TITLE OF INVENTION: Multiple Drug Resistance Gene atrd of
/ TITLE OF INVENTION: Aspergillus Nidulans
/ NUMBER OF SEQUENCES: 3
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Eli Lilly and Company
/ STREET: Lilly Corporate Center
/ CITY: Indianapolis
/ STATE: Indiana
/ COUNTRY: U.S.
/ ZIP: 46285
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/996,545
/ FILING DATE:
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Webster, Thomas D.
/ REGISTRATION NUMBER: 39,872
/ REFERENCE/DOCKET NUMBER: X-11766
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 317-276-3334
/ TELEFAX: 317-276-2763
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 4002 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: CDNA
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 1..4002
/ US-08-996-545-1

Alignment Scores:
Pred. No.: 1.56e-40 Length: 4002
Score: 366.00 Matches: 74
Percent Similarity: 73.60% Conservative: 18
Best Local Similarity: 59.20% Mismatches: 31
Query Match: 55.54% Indels: 2
DB: 2 Gaps: 1

US-09-873-409-3 (1-131) x US-08-996-545-1 (1-4002)
QY 1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
Db 3445 CTTGTTGATGGGAGGACATAAGTAAACTAAATATCACTTACCAGCAGCTTCTGTCA 3504
QY 21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
Db 3505 CTGCTCAGCAGGAGCGGACACTGTACCGGGCCACCATCAAGGAAAAACATCTTACTTGGT 3564
QY 41 -----ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAla 58
Db 3565 ATTGTCGAAGATGACGTACCGGAAGATTTCTGATTAAAGCTTGCAGGACGCTAATATC 3624
QY 59 TyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAla 78
Db 3625 TAGGACTTCATCATGCTCGCTCCGAGGGGCTTTAATACAGTTGTTGGCAAGGAGGC 3684
```


QY 79 GlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaAlaArgAlaLeuValArgAsnPro 98
Db 3685 ATGTTCTTGGGGCCCAAGCAACGTGTGGCCATTGCCGAGCCCTTCTCGGGATCCC 3744
QY 99 LysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaVal 118
Db 3745 AAAATCTTCTTCTCGATGAAGGACGTAGCCCTCGACTCCGATTCAGAAAAGGTGCTC 3804
QY 119 GlnAlaAlaLeuGlu 123
Db 3805 CAGCGGCTTTGGAT 3819
RESULT 14
US-08-996-545-3
; Sequence 3, Application US/08996545
; Patent No. 5928898
; GENERAL INFORMATION:
; APPLICANT: Skatrud, Paul L.
; APPLICANT: de Waard, Maarten A.
; APPLICANT: Peery, Robert B.
; APPLICANT: Andrade, Alan C.
; TITLE OF INVENTION: Multiple Drug Resistance Gene at rD of
; TITLE OF INVENTION: Aspergillus nidulans
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: U.S.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/08/996,545
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Webster, Thomas D.
; REGISTRATION NUMBER: 39,872
; REFERENCE/DOCKET NUMBER: X-11766
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3334
; TELEFAX: 317-276-2763
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4002 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-996-545-3
Alignment Scores:
Pred. No.: 1.56e-40 Length: 4002
Score: 366.00 Matches: 74
Percent Similarity: 73.60% Conservative: 18
Best Local Similarity: 59.20% Mismatches: 31
Query Match: 55.54% Indels: 2
DB: 2 Gaps: 1
US-09-873-409-3 (1-131) x US-08-996-545-3 (1-4002)
QY 1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
Db 3445 CUUGUUGAGGGAGACAUACAAACUAAUACUCCUCCGAGCUUUGUCA 3504
QY 21 ValValSerGlnGluProValLeuPheGlyThrIleSerAsnAsnIleTyrGly 40

Db 3505 CUGGUCAGCCAGGCGGACACUGUACCAAGGCGCACCAUCAAAGAAACAUCUUAUUGGU 3564
QY 41 -----ArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAla 58
Db 3565 AUUGUGAAGAGUAGUACCGGAGAAUUCUUAAGCUUGCAAGGACGACGCUAAUUC 3624
QY 59 TyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyAla 78
Db 3625 UACGACUUCAUCAUGUCGCCGAGGCGCUUUAUACAGUUGUUGGACGACGAGGAGGC 3684
QY 79 GlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaAlaArgAlaLeuValArgAsnPro 98
Db 3685 AUGUUGUCUGGCGGCAAAAGCAAGUGUGGCAUUGCCGAGCCCUUUCUGGGAUCCC 3744
QY 99 LysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaVal 118
Db 3745 AAAAUUU 118
QY 119 GlnAlaAlaLeuGlu 123
Db 3805 CAGCGGCUUUGGAU 3819
RESULT 15
US-09-328-320-1
; Sequence 1, Application US/09328320
; Patent No. 6228615
; GENERAL INFORMATION:
; APPLICANT: Skatrud, Paul L.
; APPLICANT: de Waard, Maarten A.
; APPLICANT: Peery, Robert B.
; APPLICANT: Andrade, Alan C.
; TITLE OF INVENTION: Multiple Drug Resistance Gene at rD of
; TITLE OF INVENTION: Aspergillus nidulans
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: U.S.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/09/328,320
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION NUMBER: 08/996,545
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Webster, Thomas D.
; REGISTRATION NUMBER: 39,872
; REFERENCE/DOCKET NUMBER: X-11766
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3334
; TELEFAX: 317-276-2763
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4002 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..4002

US-09-328-320-1

Alignment Scores:

Pred. No.:	1.56e-40	Length:	4002
Score:	366.00	Matches:	74
Percent Similarity:	73.60%	Conservative:	18
Best Local Similarity:	59.20%	Mismatches:	31
Query Match:	55.54%	Indels:	2
DB:	4	Gaps:	1

US-09-873-409-3 (1-131) x US-09-328-320-1 (1-4002)

Qy	1	MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly	20
Db	3445	CTTGTTCATGGGAAGGACATAAGTAAATATCAACTCCTACCGCAGCTTCTGTCA	3504
Qy	21	ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly	40
Db	3505	CTGGTCAGCCAGGCGGACACTGTACCGGGCACCATCAAGGAAACATCTTACTTGGT	3564
Qy	41	-----ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAla	58
Db	3565	ATTGTCGAAGATGACGTACCGGAGAAATTCTTGATTAAGCTTGCAGGACGCTAATATC	3624
Qy	59	TyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAla	78
Db	3625	TACGACTTCATCATGTGCTCCCGGAGGGCTTTAATACAGTTGTTGGCAGCAAGGAGGC	3684
Qy	79	GlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnPro	98
Db	3685	ATGTTGTCTGGCGGCCAAAGCAACGTCGTGCCATTGCCCGAGCCCTTCTTCGGGATCC	3744
Qy	99	LysIleIleuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaVal	118
Db	3745	AAATCCCTTCTTCGATGAAGCGACGTCAGCCCTCGACTCCGAGTCAGAAAGGTCGTC	3804
Qy	119	GlnAlaAlaLeuGlu	123
Db	3805	CAGGCGGCTTTGGAT	3819

Search completed: March 31, 2003, 15:23:26
Job time : 39.1843 secs

GenCore version 5.1.4 p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 30, 2003, 03:22:44 ; Search time 901.042 Seconds
(without alignments)
2354.618 Million cell updates/sec

Title: US-09-873-409-3
Perfect score: 659
Sequence: 1 MVDENDIRALNVRHYRDHIG.....SESKSAVQALEKDTPRYSF 131

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n model -DEV=xlh
-Q/cgn2_1/USPTO_spool/US9873409/runat_27032003_115420_19240/app_query.fasta_1.7544
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US9873409 @CGN 1 1 30544 @runat_27032003_115420_19240 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST.*
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_htc.*
9: gb_est1.*
10: gb_est2.*
11: gb_htc.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: gb_gss.*
18: em_gss_hum.*
19: em_gss_inv.*
20: em_gss_pln.*
21: em_gss_vrt.*
22: em_gss_fun.*
23: em_gss_mam.*
24: em_gss_mus.*
25: em_gss_oth.*
26: em_gss_pro.*
27: em_gss_rod.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	618	93.8	760	13	BM016204	BM016204 603642659
2	578	87.7	559	12	BF692596	BF692596 602248949
3	524	79.5	545	9	AL040762	AL040762 DKEZ8434C
4	490	74.4	669	13	BM013981	BM013981 603639619
5	464.5	70.5	1033	14	BM926413	BM926413 AGENCOURT
6	445	67.5	550	13	BJ075336	BJ075336 BJ075336
7	428	64.9	1098	17	CNS051EN	AL316760 Tetraodon
8	393	59.6	550	13	BJ388678	BJ388678 BJ388678
9	379	57.5	520	13	BJ336809	BJ336809 BJ336809
10	378.5	57.4	589	9	AI666987	AI666987 fc24c03.y
11	378.5	57.4	589	9	AI666988	AI666988 fc24c04.y
12	377	57.2	542	13	BJ339421	BJ339421 BJ339421
13	377	57.2	547	13	BJ333166	BJ333166 BJ333166
14	374	56.8	589	13	BJ356968	BJ356968 BJ356968
15	374	56.8	608	13	BJ356648	BJ356648 BJ356648
16	374	56.8	656	13	BJ401771	BJ401771 BJ401771
17	374	56.8	659	13	BJ350425	BJ350425 BJ350425
18	374	56.8	687	13	BJ375176	BJ375176 BJ375176
19	374	56.8	688	13	BJ405288	BJ405288 BJ405288
20	374	56.8	690	13	BJ408755	BJ408755 BJ408755
21	374	56.8	692	13	BJ407928	BJ407928 BJ407928
22	374	56.8	696	13	BJ348083	BJ348083 BJ348083
23	374	56.8	722	13	BJ406558	BJ406558 BJ406558
24	374	56.8	731	13	BJ354552	BJ354552 BJ354552
25	374	56.8	733	13	BJ376217	BJ376217 BJ376217
26	374	56.8	739	13	BJ404475	BJ404475 BJ404475
27	374	56.8	741	13	BJ347139	BJ347139 BJ347139
28	374	56.8	742	13	BJ372910	BJ372910 BJ372910
29	374	56.8	743	13	BJ408863	BJ408863 BJ408863
30	374	56.8	745	13	BJ405178	BJ405178 BJ405178
31	374	56.8	758	13	BJ357956	BJ357956 BJ357956
32	374	56.8	783	13	BJ352760	BJ352760 BJ352760
33	374	56.8	784	13	BJ405287	BJ405287 BJ405287
34	374	56.8	792	13	BJ352404	BJ352404 BJ352404
35	374	56.8	814	13	BJ352304	BJ352304 BJ352304
36	374	56.8	820	13	BJ356540	BJ356540 BJ356540
37	374	56.8	835	13	BJ349604	BJ349604 BJ349604
38	374	56.8	838	13	BJ355307	BJ355307 BJ355307
39	372	56.4	943	9	AL520322	AL520322 AL520322
40	370	56.1	681	14	BU006064	BU006064 QGG9N16.y
41	370	56.1	715	14	BQ869512	BQ869512 QGD6G09.y
42	370	56.1	894	12	BF584668	BF584668 602098406
43	369	56.0	508	14	BQ818890	BQ818890 1030074A0
44	369	56.0	635	13	BI716519	BI716519 1031010G1
45	368.5	55.9	871	9	AL582625	AL582625 AL582625

ALIGNMENTS

RESULT 1
BM016204
LOCUS 603642659F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:5418615 5',
DEFINITION mRNA sequence.
ACCESSION BM016204.1 GI:16530558
VERSION EST.
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 760)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL
COMMENT

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL2068 row: e column: 16
High quality sequence stop: 740.

FEATURES

source

1. .760
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5418615"
/clone_lib="NIH_MGC_87"
/tissue_type="mammary adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: breast; Vector: pCMV-SpORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.383 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH MGC Library."
BASE COUNT 229 a 150 c 187 g 193 t 1 others
ORIGIN

Alignment Scores:

Pred. No.: 5.63e-73 Length: 760
Score: 618.00 Matches: 128
Percent Similarity: 96.97% Conservative: 0
Best Local Similarity: 96.97% Mismatches: 2
Query Match: 93.78% Indels: 2
DB: 13 Gaps: 0

US-09-873-409-3 (1-131) x BM016204 (1-760)

QY 1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
DB 246 ATGTGGATGAGATGATCATCAGAGCTTTAAATGTGGCATTTATCGAGACCATATNGGA 305

QY 21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
DB 306 GTGCTTAGTCAAGAGCCTGTTTGTTCGGACCAACCACATCATGTAACAATATCAAGTATGGA 365

QY 41 ArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
DB 366 CGAGATGATGTGACTGATGAAGAGATGGAGAGACGACGAAGGGAAGCAATGCGTATGAT 425

QY 61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80
DB 426 TTTATCATGGATTTCTTAATAATTTAATACATTTGTTAGGGGNAAGAGCTCAATG 485

QY 81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysI 100
DB 486 AGTGGAGGCGAGAAACAGAGATCGCAATTTGGCTCGTCTAGTTCGAACCCCAAGAT 545

QY 100 eLeuIleLeu-AspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGln 120
DB 546 TCTGATTTTCAGATGAGGCTACGTCCTCCCTGGATTCCAGAAAGCAAGTCAGCTGTTCCAG 605

QY 120 laAlaLeuGluLysAspThrProArgTyrSer 130
DB 606 CTGCACCTGGAGAGGATACCCCGAGGTATTC 637

RESULT 2
BF692596
LOCUS BF692596
DEFINITION 602248949P1 NIH_MGC_62 Homo sapiens cDNA clone IMAGE:4334100 5',
mRNA sequence.
ACCESSION BF692596
VERSION BF692596.1 GI:11978004

EST 22-DEC-2000

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

human.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 559)

NIH-MGC http://mgc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC/DCTD/DTP

CDNA Library Preparation: CLONETECH Laboratories, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLNL2068 row: a column: 13

High quality sequence stop: 555.

Location/Qualifiers

1. .559

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:4334100"

/clone_lib="NIH_MGC_62"

/tissue_type="melanotic melanoma, high MDR"

/lab_host="DH10B (T1 phage-resistant)"

/notes="Organ: skin; Vector: pDNR-LJB (Clontech); Site_1:

SfiI (ggccattagggc); Site_2: SfiI (ggccattagggc);

Double-stranded cDNA was prepared from cell line RNA. 5'

and 3' adaptors were used in cloning as follows: 5'

adaptor sequence: 5'-CAGGCGCATATGGCC-3' and 3' adaptor

sequence: 5'-ATTCTAGAGCGGCGGCGGCGCATG-dt(30)BN-3'

(where B = A, C, or G and N = A, C, G, or T). Average

insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies

contained inserts by PCR. This library was enriched for

full-length clones and was constructed by Clontech

Laboratories (Palo Alto, CA)."

BASE COUNT 161 a 118 c 141 g 139 t

ORIGIN

Alignment Scores:

Pred. No.: 9.32e-68 Length: 559
Score: 578.00 Matches: 126
Percent Similarity: 96.95% Conservative: 1
Best Local Similarity: 96.18% Mismatches: 3
Query Match: 87.71% Indels: 3
DB: 12 Gaps: 0

US-09-873-409-3 (1-131) x BF692596 (1-559)

QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
DB 28 GTGCGATGAGAATGACATCAGAGCTTTAAATGTGCGCATTTATCGAGACCATAT-GGAGTG 86

QY 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
DB 87 GTTAGTCAGAGAGCTGT-TGTTCGGGACCAACCATCAGTAACATATCAAGTATGACGA 145

QY 42 AspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
DB 146 GATGATGTGACTGATGAAGAGATGGAGAGACGACGAAGGAAGCAATGATGATGTTT 205

QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLys-GlyAlaGlnMetSe 81
DB 206 ATCATGAGTTTCTTAATAATTTAATACATTGTTAGGGGAAACAGGAGCTCAATGAG 265

QY 81 rGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLe 101
DB 266 TGGAGGGCAGAACACAGAGATCGCAATTCCTGCTCTAGTTCGAACCCCAAGATTCT 325

QY 101 uIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaA 121


```

Db 326 GATTTTAGATGAGGCTACCTCTGCGCCCTGATTCAGAAACGAGTCAGCTGTTCAAGCTGC 385
|||||
Qy 121 aleuGluLysAspThrProArgTyrSerPhe 131
|||||
Db 386 ACTGGAGAGGATACCCCGAGGTATTCATT 416
|||||

RESULT 3
AL040762 545 bp mRNA linear EST 29-FEB-2000
LOCUS DKFZP434C1815 r1 434 (synonym: htes3) Homo sapiens cDNA clone
DEFINITION DKFZP434C1815 5', mRNA sequence.
ACCESSION AL040762
VERSION AL040762.1 GI:5409708
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 545)
AUTHORS Bloecker,H., Boecker,M., Brandt,P., Mewes,H.W., Gassenhuber,J. and
Wiemann,S.
TITLE EST (Bloecker, et al.)
JOURNAL Unpublished (1999)
COMMENT Contact: Bloecker H
MIPS
Am Klopferseitz 18a D-82152 Martinsried, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by GBF (National Research Centre for Biotechnology Ltd.,
Braunschweig/Germany) within the cDNA sequencing consortium of the
German Genome Project.
s1 sequence also available.
This clone (DKFZP434C1815) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
source
1..545
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZP434C1815"
/clone_lib="434 (synonym: htes3)"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/notes="Vector: pSport1; Site_1: NotI; Site_2: SalI"
BASE COUNT 171 a 89 c 138 g 147 t
ORIGIN
Alignment Scores:
Pred. No.: 1,76e-60 Length: 545
Score: 524.00 Matches: 103
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 79.51% Indels: 0
DB: 9 Gaps: 0

US-09-873-409-3 (1-131) x AL040762 (1-545)
Qy 1 MetValAspGluAsnAspLeuArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
Db 237 ATGGTGGATGAGAAATGACATCAGAGCTTTAAATGTGGCGCATTCAGAGACCATATTGGA 296
Qy 21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
Db 297 GTGGTTAGTCAAGAGCGCTGTTTGTTCGGACCAACCATCAGTAGTAACAATATCAAGTAGTGA 356
Qy 41 ArgAspAspValThrAspGluGluMetGluArgAlaLeuArgGluAlaAsnAlaTyrAsp 60
Db 357 CGAGATGATGTGCTGATGAAGAGATGGAGAGCAGCAGGAGGAGCAATATGCGTATGAT 416
Qy 61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80

```

```

Db 417 TTATCATGGAGTTTCCATAATAATTTAATCATTTGGTAGGGAAAAAGAGCTCAATG 476
|||||
Qy 81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100
|||||
Db 477 ACTGGAGGGCAGAAACAGAGGATCGCAATTGCTCGTCTAGTTCGAAACCCCAAGATT 536
|||||
Qy 101 LeuIleLeu 103
|||||
Db 537 CTGATTTTA 545
|||||

RESULT 4
EM013981
LOCUS 603639619F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:5415556 5',
DEFINITION mRNA sequence.
ACCESSION EM013981
VERSION EM013981.1 GI:16528335
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 669)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLN12060 row: f column: 05
High quality sequence stop: 664.
Location/Qualifiers
1..669
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5415556"
/clone_lib="NIH_MGC_87"
/tissue_type="mammary adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: breast; Vector: pCMV-SPORT6; Site 1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.383 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT 203 a 122 c 162 g 182 t
ORIGIN
Alignment Scores:
Pred. No.: 9.06e-56 Length: 669
Score: 490.00 Matches: 98
Percent Similarity: 98.99% Conservative: 0
Best Local Similarity: 98.99% Mismatches: 0
Query Match: 74.36% Indels: 1
DB: 13 Gaps: 0

US-09-873-409-3 (1-131) x EM013981 (1-669)
Qy 1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
Db 374 ATGGTGGATGAGAAATGACATCAGAGCTTTAAATGTGGCGCATTCAGAGACCATATTGGA 433
Qy 21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
Db 434 GTGGTTAGTCAAGAGCGCTGTTTGTTCGGACCAACCATCAGTAGTAACAATATCAAGTAGTGA 493
Qy 41 ArgAspAspValThrAspGluGluMetGluArgAlaLeuArgGluAlaAsnAlaTyrAsp 60

```



```

Db 494 CGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCGTATGAT 553
Qy 61 PheileMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80
Db 554 TTTTATCATGAGATTTCTTAATAAATTTAATACATTTGGTAGGGGAAAAGAGAGCTCAATG 613
Qy 81 SerGlyGlyGlnLysGlnArg-1leAlaileAlaArgAlaLeuValArgAsnPro 98
Db 614 AGTGGAGGCGAGAAACAGAGGCATCGCAATTTGCTGTGCTTATGTTTGAACCC 668

RESULT 5
BM926413
LOCUS BM926413
DEFINITION AGENCOURT_6600788 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5764845
5', mRNA sequence.
ACCESSION BM926413
VERSION BM926413.1 GI:19376792
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1033)
NIH-MGC http://img.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapps@email.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12818 row: o column: 22
High quality sequence start: 3
High quality sequence stop: 686.
Location/Qualifiers
1. .1033
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5764845"
/clone_lib="NIH_MGC_114"
/lab_host="DH10B"
/notes="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: EcoRV (destroyed); RNA source anonymous pool of
male brains, age range 23-27 yo. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.5 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(invitrogen). Research Genetics tracking code 019. Note:
this is a NIH_MGC Library."
BASE COUNT 289 a 212 c 218 g 312 t 2 others
ORIGIN
source

Alignment Scores:
Pred. No.: 4,54e-52 Length: 1033
Score: 464.50 Matches: 111
Percent Similarity: 87.02% Conservative: 3
Best Local Similarity: 84.73% Mismatches: 17
Query Match: 70.49% Indels: 4
DB: 14 Gaps: 0

US-09-873-409-3 (1-131) x BM926413 (1-1033)
Qy 1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
Db 534 ATGGTGGATGAGAAATGACATCAGACGCTTTAAATGTGGCGCATATATGACCAATATGGA 593
Qy 21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40

```

```

Db 594 GTGGTTAGTCAAGAGCCTGTTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGA 653
Qy 41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
Db 654 CGAGATGATGTGCTGACTGATGAAGAGATGGAGAGCAGCAGCAAGGGAAGCAAATGCGTATGAT 713
Qy 61 PheileMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80
Db 714 TTTTATCATGAGATTTCTTAATAA-TTTAATACATTTGGTAGGGGAAAAGAGAGCTCAATG 772
Qy 81 SerGlyGlyGlnLysGlnArgIleAlaileAlaArgAlaLeuValArgAsnProLysIle 100
Db 773 AGTGGAGGCGAGAAACCG-AGGATCCCAATTTGCTGCTCTAGTTTCGAAACCCAGGAT- 830
Qy 101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120
Db 831 CTGAATTTAGATGAAGCTACGTCTGCCCTGGATTTCAGAAACCAAGTCAACTGGTCA--GCT 888
Qy 121 AlaLeuGluLysAspThrProArgTyrSerPhe 131
Db 889 GGCTGGAGAGGAAACCCCGGNAATTCCTTT 921

RESULT 6
BM925336
LOCUS BM925336
DEFINITION laevis cDNA clone XL057d10 5', mRNA sequence.
ACCESSION BM925336
VERSION BM925336.1 GI:17520252
KEYWORDS EST.
SOURCE African clawed frog.
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Xenopus.
1 (bases 1 to 550)
REFERENCE 1
AUTHORS Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and Kohara
,Y.
TITLE Expressed genes in X. laevis embryo
JOURNAL Unpublished (2001)
COMMENT Contact: Tadaeu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
1. .550
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/clone="XL057d10"
/clone_lib="NIBB Mochii normalized Xenopus tailbud
library"
/tissue type="whole embryo"
/dev stage="stage 25"
/notes="Vector: PBSRN3; Site_1: NotI; Site_2: EcoRI; cDNAs
were oligo-dT primed and directionally cloned. Staging
according to Nieuwkoop and Faber. Library is subtracted
and was constructed by N. Garrett and A.M. Zorn,
(Wellcome/CRC Institute)."
BASE COUNT 162 a 109 c 144 g 135 t
ORIGIN
source

Alignment Scores:
Pred. No.: 8,26e-50 Length: 550
Score: 445.00 Matches: 86
Percent Similarity: 90.00% Conservative: 22
Best Local Similarity: 71.67% Mismatches: 12
Query Match: 67.53% Indels: 0
DB: 13 Gaps: 0

```



```

/clone="ddse6p15"
/clone_lib="Dictyostelium discoideum cDNA library, SP"
/sex="mat A"
/dev_stage="Slug stage"
BASE COUNT 190 a -77 c 113 g 170 t
ORIGIN

```

```

Alignment Scores:
Pred. No.: 8.66e-43 Length: 550
Score: 393.00 Matches: 76
Percent Similarity: 78.86% Conservatives: 21
Best Local Similarity: 61.79% Mismatches: 26
Query Match: 59.64% Indels: 0
DB: 13 Gaps: 0

```

US-09-873-409-3 (1-131) x BJ388678 (1-550)

```

QY 1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
DB 96 TTATTGGATGGAGAGATATTAGAAAATTCATATGTCAGAGGGTTACGTCAAAAAATTTGTT 155

```

```

QY 21 ValValSerGlnGluProValLeuPheGlyThrIleSerAsnAsnIleLysTyrGly 40
DB 156 TTATGTCATCAAGAGCGCTGTATTATTCGCCCAACCATTAGTGAATAATTTCTGTTATGTT 215

```

```

QY 41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
DB 216 AAGAAGGTGCCACTCAAGATGAATTAAGAAGCAGCTAAATAGCAATGCCATAGT 275

```

```

QY 61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80
DB 276 TTTATATCACAATTACACACAGGGTTATATATCTTTGTTAGTGCAGAAAGGTGCCAAATG 335

```

```

QY 81 SerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100
DB 336 TCTGGTGGTCAAGACACACGTATAGCAATTCGTCGCGCTTTATTAATAATCCAAATATA 395

```

```

QY 101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120
DB 396 TTACTCTTGATGATCAACATCTGCGATGGATGCAGAGTCTACAAATTAGTACAAGAG 455

```

```

QY 121 AlaLeuGlu 123
DB 456 GCATTAGAT 464

```

```

RESULT 9
BJ336809
LOCUS
DEFINITION
Dictyostelium discoideum cDNA library, AF Dictyostelium
discoideum cDNA clone dda55p03 5', mRNA sequence.
ACCESSION
VERSION
SOURCE
ORGANISM
Eukaryota; Metazoa; Dictyosteliida; Dictyostelium.

```

```

REFERENCE
1 (bases 1 to 520)
Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.
Full length cDNA of Dictyostelium discoideum at the aggregation
stage
Unpublished (2002)
Contact: Tadao Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
1. .520
/organism="Dictyostelium discoideum"
/strain="AX4"
/db_xref="taxon:44689"

```

JOURNAL

COMMENT

```

Contact: Stephen L. Johnson
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: zbrafish@watson.wustl.edu
CDNA Library Preparation: Matthew Clark. cDNA Library Arrayed by:
Matthew Clark. DNA Sequencing by: Washington University Genome
Sequencing Center Clone distribution: Genome Systems, St. Louis,

```

FEATURES

source

```

/clone="dda55p03"
/clone_lib="Dictyostelium discoideum cDNA library, AF"
/sex="mat A"
/dev_stage="Aggregation stage"
BASE COUNT 178 a -78 c 106 g 158 t
ORIGIN

```

```

Alignment Scores:
Pred. No.: 6.23e-41 Length: 520
Score: 379.00 Matches: 73
Percent Similarity: 78.81% Conservatives: 20
Best Local Similarity: 61.86% Mismatches: 25
Query Match: 57.51% Indels: 0
DB: 13 Gaps: 0

```

US-09-873-409-3 (1-131) x BJ336809 (1-520)

```

QY 1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
DB 165 TTATTGGATGGAGAGATATTAGAAAATTCATATGTCAGAGGGTTACGTCAAAAAATTTGTT 224

```

```

QY 21 ValValSerGlnGluProValLeuPheGlyThrIleSerAsnAsnIleLysTyrGly 40
DB 225 TTATGTCATCAAGAGCGCTGTATTATTCGCCCAACCATTAGTGAATAATTTCTGTTATGTT 284

```

```

QY 41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
DB 285 AAGAAGGTGCCACTCAAGATGAATTAAGAAGCAGCTAAATAGCAATGCCATAGT 344

```

```

QY 61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80
DB 345 TTTATATCACAATTACACACAGGGTTATATATCTTTGTTAGTGCAGAAAGGTGCCAAATG 404

```

```

QY 81 SerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100
DB 405 TCTGGTGGTCAAGACACACGTATAGCAATTCGTCGCGCTTTATTAATAATCCAAATATA 464

```

```

QY 101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaVal 118
DB 465 TTACTCTTGATGATCAACATCTGCGATGGATGCAGAGTCTACAAATTAGTGA 518

```

```

RESULT 10
A1666987
LOCUS
DEFINITION

```

```

A1666987 588 bp mRNA linear EST 07-JUN-2001
fc24c03.y1 Zebrafish WashU MPIMG EST Danio rerio cDNA clone
IMAGE:3722308 5' similar to TR:Q13040 Q13040 ATP-BINDING CASSETTE
PROTEIN ;, mRNA sequence.
A1666987.1 GI:4805343
VERSION
KEYWORDS
SOURCE
ORGANISM
Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
; Cyprinidae; Danio.
1 (bases 1 to 588)
Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy
,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood
,K., Steptoe,L., Theising,B., Allen,M., Bowers,Y., Person,B.,
Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Richter,E.,
Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R.
and Wilson,R.
WashU Zebrafish EST Project 1998
Unpublished (1998)
Contact: Stephen L. Johnson
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: zbrafish@watson.wustl.edu
CDNA Library Preparation: Matthew Clark. cDNA Library Arrayed by:
Matthew Clark. DNA Sequencing by: Washington University Genome
Sequencing Center Clone distribution: Genome Systems, St. Louis,

```

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes

; Cyprinidae; Danio.

1 (bases 1 to 588)

Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy

,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood

,K., Steptoe,L., Theising,B., Allen,M., Bowers,Y., Person,B.,

Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Richter,E.,

Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R.

and Wilson,R.

WashU Zebrafish EST Project 1998

Unpublished (1998)

Contact: Stephen L. Johnson

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: zbrafish@watson.wustl.edu

CDNA Library Preparation: Matthew Clark. cDNA Library Arrayed by:

Matthew Clark. DNA Sequencing by: Washington University Genome

Sequencing Center Clone distribution: Genome Systems, St. Louis,

Missouri (web address: www.genomesystems.com) (email contact: info@genomesystems.com) and Research Genetics, Huntsville, Alabama (web address: www.resgen.com) (email contact: info@resgen.com) and RessourcenzentrumPrimatDatenbank, Berlin, Germany (web address: www.rzpd.de)

Seq primer: T3 ET from Amersham
High quality sequence stop: 481.
Location/Qualifiers

FEATURES

source

1..588
/organism="Danio rerio"
/db_xref="taxon:7955"
/clone="IMAGE:3722308"
/clone_lib="Zebrafish Washu MPING EST"
/sex="mixed"
/tissue type="26 somite embryos, adult livers, shield stage embryos"
/lab host="XLI-blue MRF"
/note="Vector: pSPORT1; Site 1: NotI; Site 2: SalI; 1st strand cDNA was primed with a Not I - oligo(dT)15 primer 5'-pGACTAGTCTAGTCGAGCGCGCCCTTTTCTTTTCTTTT3'; double-stranded cDNA was ligated to Sal I adaptors (BRL), digested with Not I and cloned into the Not I and Sal I sites of the pSPORT1 vector (BRL). Library was constructed by Matthew Clark (Lehrach lab; ICRF, London and Max Planck Institut fuer Molekulare Genetik, Berlin). cDNAs for EST analysis were selected following oligonucleotide hybridization fingerprinting of arrayed clones from zebrafish late somitogenesis (26 ss), adult liver or embryonic shield stage (5.6 h) libraries. Fingerprint data were used to computationally cluster cDNAs, and a single cDNA from each cluster was chosen for sequencing. In some cases multiple members of the same cluster were sequenced to assess clustering parameters or single clones were sequenced additional times to assess quality control."

BASE COUNT 149 a 138 c 167 g 132 t 2 others
ORIGIN

Alignment Scores:
Pred. No.: 8.61e-41 Length: 588
Score: 378.50 Matches: 79
Percent Similarity: 77.78% Conservative: 19
Best Local Similarity: 62.70% Mismatches: 25
Query Match: 57.44% Indels: 3
DB: 9 Gaps: 1

US-09-873-409-3 (1-131) x AI666987 (1-588)

Qy 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
Db 1 GTTGATGGTCATGATATTCGAGACTTGAATCCCTTACTGGTTAGGAGTCACATTCGTACT 60
Qy 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleValTyrGlyArg 41
Db 61 GTAAGCCAGGAACCGGTCTCTTCTCTGTCGTTCAATTCGTGAGAACATTCGTTATGAGCG 120
Qy 42 AspAsp-----ValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAla 58
Db 121 CCGATCCGAGTCAGTGACAGTCGAGGAGACATTTACAGCGCGCACAGATCGCAATGCT 180
Qy 59 TyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyValAla 78
Db 181 CATGACTTCATCCAGGAGTTTCCCAAGGGGTACGACTACTGTGCTGGGAGAGAGGGGGTT 240
Qy 79 GlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnPro 98
Db 241 CTGCTGTCTAGGTGGAACAAGACAGAGAGTGGCCATGACGACGACCTGCTGGAAGATCCC 300
Qy 99 LysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaVal 118
Db 301 AAATACTTTTACTGATGAGCCAGCCAGTGCTCTGTGATGACAGATGATGATGTTTGGTC 360
Qy 119 GlnAlaLeuGlyLys 124

Db 361 CAGGAGCGCTTAGAAAGG 378

RESULT 11

AI666988

LOCUS

DEFINITION

fc24c04.y1 Zebrafish Washu MPING EST Danio rerio cDNA clone

IMAGE:3722310 5' similar to TR:Q13040 Q13040 ATP-BINDING CASSETTE

PROTEIN ; mRNA sequence.

AI666988

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy

,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood

,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B.,

Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E.,

Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R.

and Wilson,R.

Washu Zebrafish EST Project 1998

Unpublished (1998)

Contact: Stephen L. Johnson

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: zbrafish@watson.wustl.edu

CDNA Library Preparation: Matthew Clark. cDNA Library Arrayed by:

Matthew Clark. DNA Sequencing by: Washington University Genome

Sequencing Center Clone distribution: Genome Systems, St. Louis,

Missouri (web address: www.genomesystems.com) (email contact:

info@genomesystems.com) and Research Genetics, Huntsville, Alabama

(web address: www.resgen.com) (email contact: info@resgen.com) and

RessourcenzentrumPrimatDatenbank, Berlin, Germany (web address:

www.rzpd.de)

Seq primer: T3 ET from Amersham

High quality sequence stop: 481.

Location/Qualifiers

1..588

/organism="Danio rerio"

/db_xref="taxon:7955"

/clone="IMAGE:3722310"

/clone_lib="Zebrafish Washu MPING EST"

/sex="mixed"

/tissue type="26 somite embryos, adult livers, shield

stage embryos"

/lab host="XLI-blue MRF"

/note="Vector: pSPORT1; Site 1: NotI; Site 2: SalI; 1st

strand cDNA was primed with a Not I - oligo(dT)15 primer

5'-pGACTAGTCTAGTCGAGCGCGCCCTTTTCTTTTCTTTT3';

double-stranded cDNA was ligated to Sal I adaptors (BRL),

digested with Not I and cloned into the Not I and Sal I

sites of the pSPORT1 vector (BRL). Library was constructed

by Matthew Clark (Lehrach lab; ICRF, London and Max Planck

Institut fuer Molekulare Genetik, Berlin). cDNAs for EST

analysis were selected following oligonucleotide

hybridization fingerprinting of arrayed clones from

zebrafish late somitogenesis (26 ss), adult liver or

embryonic shield stage (5.6 h) libraries. Fingerprint

data were used to computationally cluster cDNAs, and a

single cDNA from each cluster was chosen for sequencing.

In some cases multiple members of the same cluster were

sequenced to assess clustering parameters or single clones

were sequenced additional times to assess quality

control."

BASE COUNT 151 a 137 c 167 g 133 t 1 others

ORIGIN

Alignment Scores:

Pred. No.: 8.63e-41 Length: 589
 Score: 378.50 Matches: 79
 Percent Similarity: 77.78% Conservative: 19
 Best Local Similarity: 62.70% Mismatches: 25
 Query Match: 57.44% Indels: 3
 DB: 9 Gaps: 1

US-09-873-409-3 (1-131) x AI666988 (1-589)

QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
 DB 1 GTTGATGGTCTGATATTCGAGACTTGAATCCTTATTGGTTAAGGAGTCACATGGTACT 60
 QY 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
 DB 61 GTAAGCCAGAACCGGTCCTTTCTCTGTTCATATGCTCAGAACATGCTTATGAGCG 120
 QY 42 AspAsp-----ValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAla 58
 DB 121 CCGATCCGAGTCAGGTGACAGTGGAGGACATTTACAGGCGGCGCACAGATCGCCATGCT 180
 QY 59 TyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyVala 78
 DB 181 CATGACTTCATCCAGGAGTTTCCCAAGGGCTACGATACCTGTGTGGAGAGAGGGGGT 240
 QY 79 GlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnPro 98
 DB 241 CTGCTGTTCAGTGGACAAAGACGAGAGTGGCCATAGCACAGCCCTGCTGAAGATCCC 300
 QY 99 LysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaVal 118
 DB 301 AAAATACTTTTACTGGATGAAGCCAGGCTCTGGATGCAGAGAATGAGTTTGGTC 360
 QY 119 GlnAlaLeuGluLys 124
 DB 361 CAGGAGCGGTAGAAAGG 378

RESULT 12

BJ339421

LOCUS

BJ339421 Dictyostelium discoideum cDNA library, AF Dictyostelium
 discoideum cDNA clone dda65d11 5', mRNA sequence.

ACCESSION

BJ339421

VERSION

BJ339421.1 GI:19247783

KEYWORDS

Dictyostelium discoideum.

SOURCE

Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.

REFERENCE

1 (bases 1 to 542)

AUTHORS

Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.

TITLE

Full length cDNA of Dictyostelium discoideum at the aggregation

JOURNAL

Unpublished (2002)

COMMENT

Contact: Tadasu Shin-i
 Center For Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshini@genes.nig.ac.jp.

FEATURES

source

1..542
 /organism="Dictyostelium discoideum"
 /strain="AX4"
 /db_xref="taxon:44689"
 /clone="dda65d11"
 /sex="mat A"
 /dev_stage="Aggregation stage"

BASE COUNT

186 a

77 c

113 g

166 t

ORIGIN

Alignment Scores:

Pred. No.: 1.23e-40 Length: 542
 Score: 377.00 Matches: 72
 Percent Similarity: 80.70% Conservative: 20
 Best Local Similarity: 63.16% Mismatches: 22
 Query Match: 57.21% Indels: 0
 DB: 13 Gaps: 0

US-09-873-409-3 (1-131) x BJ339421 (1-542)

QY 1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
 DB 192 TTATTGGATGGAGAGATATTAGAAAATTCATGTGAGAGGTTAGCTCAAAAATTTGGT 251
 QY 21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
 DB 252 TTATGTCATCAAGAGCGCTGTATTATTCGCCACACACCATTTAGTGAATAATTCGTTATGGT 311
 QY 41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
 DB 312 AAAGAAGGTGCCACTCAGATGAAATTTGAAGAAGCAGCTAAATTTAGCAATGCCATAGT 371
 QY 61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80
 DB 372 TTATATCAATCAATTCACACAGGTTATAATCTTTGTTAGGTGAGAAAGGTGTCCAAATG 431
 QY 81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100
 DB 432 TCTGGTGGTCAAGACACAGTATAGCAATTCGTGCGCGTTATTAAAAATCCAAATATA 491
 QY 101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 114
 DB 492 TTACTCTTGATGAATCAACATCTGCAATTCGATTCGATGCAGAGTCT 533

RESULT 13

BJ333166

LOCUS

BJ333166 Dictyostelium discoideum cDNA library, AF Dictyostelium
 discoideum cDNA clone dda41m18 5', mRNA sequence.

ACCESSION

BJ333166

VERSION

BJ333166.1 GI:19163296

KEYWORDS

Dictyostelium discoideum.

SOURCE

Dictyostelium discoideum

ORGANISM

Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.

REFERENCE

1 (bases 1 to 547)

AUTHORS

Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.

TITLE

Full length cDNA of Dictyostelium discoideum at the aggregation

JOURNAL

Unpublished (2002)

COMMENT

Contact: Tadasu Shin-i

Center For Genetic Resource Information

National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tshini@genes.nig.ac.jp.

FEATURES

source

1..547

/organism="Dictyostelium discoideum"

/strain="AX4"

/db_xref="taxon:44689"

/clone="dda41m18"

/sex="mat A"

/dev_stage="Aggregation stage"

186 a

78 c

114 g

169 t

Alignment Scores:

Pred. No.: 1.24e-40 Length: 547

Score: 377.00 Matches: 72

Percent Similarity: 80.70% Conservative: 20

Best Local Similarity: 63.16%	Mismatches: 22
Query Match: 57.21%	Indels: 0
DB: 13	Gaps: 0
US-09-873-409-3 (1-131) x BJ333166 (1-547)	
QY 1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyArgAspHisIleGly 20	
DB 195 TTATTGGATGGAGAAAGATATTAGAAAATTCAATGTGCAGAGGGTTAGTCAAAAATTTGGT 254	
QY 21 ValValSerGlnGluProValLeuPheGlyThrIleSerAsnAsnIleLysTyGly 40	
DB 255 TTATGTAATCAAGAGCTGTATTATTTGCCACACCAATTAGTCAAAATATTGTTATGTT 314	
QY 41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyArg 60	
DB 315 AAAGAAGGTGCCACTCAAGATGAATTAAGAGAGCAGCTAAATTAGCAAAATGCCATAGT 374	
QY 61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80	
DB 375 TTATATATCAAAATTACACAGGGTTATAATACTTTGTAGTGAGAAAGGTGTCCAAATG 434	
QY 81 SerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100	
DB 435 TCTGGTGGTCAAGAACAAGTATAGCAATTCGTGCGCGCTATTAAATAATCCAAATATA 494	
QY 101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 114	
DB 495 TTACTCTGGATGAATCAACATCTGCATTGGATGCAGAGTCT 536	
RESULT 14	
BJ356968/c	
LOCUS	
DEFINITION	BJ356968 Dictyostelium discoideum cDNA library, AF Dictyostelium
ACCESSION	BJ356968
VERSION	BJ356968.1 GI:19256563
KEYWORDS	EST.
SOURCE	Dictyostelium discoideum.
ORGANISM	Dictyostelium discoideum
REFERENCE	Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
AUTHORS	1 (bases 1 to 589)
TITLE	Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T. Full length cDNA of Dictyostelium discoideum at the aggregation stage
JOURNAL	Unpublished (2002)
COMMENT	Contact: Tadaasu Shin-i Center For Genetic Resource Information National Institute of Genetics 1111 Yata, Mishima, Shizuoka 411-8540, Japan Tel: 81-559-81-6856 Fax: 81-559-81-6855 Email: tshini@genes.nig.ac.jp. Location/Qualifiers
FEATURES	source
1..589	
/organism="Dictyostelium discoideum"	
/strain="AX4"	
/db_xref="taxon:44689"	
/clone="ddae2e02"	
/clone_lib="Dictyostelium discoideum cDNA library, AF"	
/sex="mat A"	
/dev_stage="Aggregation stage"	
BASE COUNT	182 a 103 c 99 g 205 t
ORIGIN	
Alignment Scores:	
Pred. No.:	3.5e-40
Score:	374.00
Percent Similarity:	79.51%
Best Local Similarity:	58.20%
Query Match:	56.75%
Indels:	0
Gaps:	13
Matches:	71
Conservative:	26
Mismatches:	25
Indels:	0
Gaps:	0
BASE COUNT	187 a 110 c 101 g 210 t
ORIGIN	
Alignment Scores:	
Pred. No.:	3.65e-40
Score:	374.00
Percent Similarity:	79.51%
Best Local Similarity:	58.20%
Query Match:	56.75%
Indels:	0
Gaps:	13
Matches:	71
Conservative:	26
Mismatches:	25
Indels:	0
Gaps:	0
BASE COUNT	187 a 110 c 101 g 210 t
ORIGIN	
Alignment Scores:	
Pred. No.:	3.65e-40
Score:	374.00
Percent Similarity:	79.51%
Best Local Similarity:	58.20%
Query Match:	56.75%
Indels:	0
Gaps:	13
Matches:	71
Conservative:	26
Mismatches:	25
Indels:	0
Gaps:	0
BASE COUNT	187 a 110 c 101 g 210 t
ORIGIN	
Alignment Scores:	
Pred. No.:	3.65e-40
Score:	374.00
Percent Similarity:	79.51%
Best Local Similarity:	58.20%
Query Match:	56.75%
Indels:	0
Gaps:	13
Matches:	71
Conservative:	26
Mismatches:	25
Indels:	0
Gaps:	0
BASE COUNT	187 a 110 c 101 g 210 t
ORIGIN	
Alignment Scores:	
Pred. No.:	3.65e-40
Score:	374.00
Percent Similarity:	79.51%
Best Local Similarity:	58.20%
Query Match:	56.75%
Indels:	0
Gaps:	13
Matches:	71
Conservative:	26
Mismatches:	25
Indels:	0
Gaps:	0
BASE COUNT	187 a 110 c 101 g 210 t
ORIGIN	
Alignment Scores:	
Pred. No.:	3.65e-40
Score:	374.00
Percent Similarity:	79.51%
Best Local Similarity:	58.20%
Query Match:	56.75%
Indels:	0
Gaps:	13
Matches:	71
Conservative:	26
Mismatches:	25
Indels:	0
Gaps:	0
BASE COUNT	187 a 110 c 101 g 210 t
ORIGIN	
Alignment Scores:	
Pred. No.:	3.65e-40
Score:	374.00
Percent Similarity:	79.51%
Best Local Similarity:	58.20%
Query Match:	56.75%
Indels:	0
Gaps:	13
Matches:	71
Conservative:	26
Mismatches:	25
Indels:	0
Gaps:	0
BASE COUNT	187 a 110 c 101 g 210 t
ORIGIN	
Alignment Scores:	
Pred. No.:	3.65e-40
Score:	374.00
Percent Similarity:	79.51%
Best Local Similarity:	58.20%
Query Match:	56.75%
Indels:	0
Gaps:	13
Matches:	71
Conservative:	26
Mismatches:	25
Indels:	0
Gaps:	0
BASE COUNT	187 a 110 c 101 g 210 t
ORIGIN	
Alignment Scores:	
Pred. No.:	3.65e-40
Score:	374.00
Percent Similarity:	79.51%
Best Local Similarity:	58.20%
Query Match:	56.75%
Indels:	0
Gaps:	13
Matches:	71
Conservative:	26
Mismatches:	25
Indels:	0
Gaps:	0
BASE COUNT	187 a 110 c 101 g 210 t
ORIGIN	
Alignment Scores:	
Pred. No.:	3.65e-40
Score:	374.00
Percent Similarity:	79.51%
Best Local Similarity:	58.20%
Query Match:	56.75%
Indels:	0
Gaps:	13
Matches:	71
Conservative:	26
Mismatches:	25
Indels:	0
Gaps:	0
BASE COUNT	187 a 110 c 101 g 210 t
ORIGIN	
Alignment Scores:	
Pred. No.:	3.65e-40
Score:	374.00
Percent Similarity:	79.51%
Best Local Similarity:	58.20%
Query Match:	56.75%
Indels:	0
Gaps:	13
Matches:	71
Conservative:	26
Mismatches:	25
Indels:	0
Gaps:	0
BASE COUNT	187 a 110 c 101 g 210 t
ORIGIN	
Alignment Scores:	
Pred. No.:	3.65e-40
Score:	374.00
Percent Similarity:	79.51%
Best Local Similarity:	58.20%
Query Match:	56.75%
Indels:	0
Gaps:	13
Matches:	71
Conservative:	26
Mismatches:	25
Indels:	0
Gaps:	0
BASE COUNT	187 a 110 c 101 g 210 t
ORIGIN	
Alignment Scores:	
Pred. No.:	3.65e-40
Score:	374.00
Percent Similarity:	79.51%
Best Local Similarity:	58.20%
Query Match:	56.75%
Indels:	0
Gaps:	13
Matches:	71
Conservative:	26
Mismatches:	25
Indels:	0
Gaps:	0
BASE COUNT	187 a 110 c 101 g 210 t
ORIGIN	
Alignment Scores:	
Pred. No.:	3.65e-40
Score:	374.00
Percent Similarity:	79.51%
Best Local Similarity:	58.20%
Query Match:	56.75%
Indels:	0
Gaps:	13
Matches:	71
Conservative:	26
Mismatches:	25
Indels:	0
Gaps:	0
BASE COUNT	187 a 110 c 101 g 210 t
ORIGIN	
Alignment Scores:	
Pred. No.:	3.65e-40
Score:	374.00
Percent Similarity:	79.51%
Best Local Similarity:	58.20%
Query Match:	56.75%
Indels:	0
Gaps:	13
Matches:	71
Conservative:	26
Mismatches:	25
Indels:	0
Gaps:	0
BASE COUNT	187 a 110 c 101 g 210 t
ORIGIN	
Alignment Scores:	
Pred. No.:	3.65e-40
Score:	374.00
Percent Similarity:	79.51%
Best Local Similarity:	58.20%
Query Match:	56.75%
Indels:	0
Gaps:	13
Matches:	71
Conservative:	26
Mismatches:	25
Indels:	0
Gaps:	0
BASE COUNT	187 a 110 c 101 g 210 t
ORIGIN	
Alignment Scores:	
Pred. No.:	3.65e-40
Score:	374.00
Percent Similarity:	79.51%
Best Local Similarity:	58.20%
Query Match:	56.75%
Indels:	0
Gaps:	13
Matches:	71
Conservative:	26
Mismatches:	25
Indels:	0
Gaps:	0
BASE COUNT	187 a 110 c 101 g 210 t
ORIGIN	
Alignment Scores:	
Pred. No.:	3.65e-40
Score:	374.00
Percent Similarity:	79.51%
Best Local Similarity:	58.20%
Query Match:	56.75%
Indels:	0
Gaps:	13
Matches:	71
Conservative:	26
Mismatches:	25
Indels:	0
Gaps:	0
BASE COUNT	187 a 110 c 101 g 210 t
ORIGIN	
Alignment Scores:	
Pred. No.:	3.65e-40
Score:	374.00
Percent Similarity:	79.51%
Best Local Similarity:	58.20%
Query Match:	56.75%
Indels:	0
Gaps:	13
Matches:	71
Conservative:	26
Mismatches:	25
Indels:	0
Gaps:	0
BASE COUNT	187 a 110 c 101 g 210 t
ORIGIN	
Alignment Scores:	
Pred. No.:	3.65e-40
Score:	374.00
Percent Similarity:	79.51%
Best Local Similarity:	58.20%
Query Match:	56.75%
Indels:	0
Gaps:	13
Matches:	71
Conservative:	26
Mismatches:	25
Indels:	0
Gaps:	0
BASE COUNT	187 a 110 c 101 g 210 t
ORIGIN	
Alignment Scores:	
Pred. No.:	3.65e-40
Score:	374.00
Percent Similarity:	79.51%
Best Local Similarity:	58.20%
Query Match:	56.75%
Indels:	0
Gaps:	13
Matches:	71
Conservative:	26
Mismatches:	25
Indels:	0
Gaps:	0
BASE COUNT	187 a 110 c 101 g 210 t
ORIGIN	
Alignment Scores:	
Pred. No.:	3.65e-40
Score:	374.00
Percent Similarity:	79.51%
Best Local Similarity:	58.20%
Query Match:	56.75%
Indels:	0
Gaps:	13
Matches:	71
Conservative:	26
Mismatches:	25
Indels:	0
Gaps:	0
BASE COUNT	187 a 110 c 101 g 210 t
ORIGIN	
Alignment Scores:	
Pred. No.:	3.65e-40
Score:	374.00
Percent Similarity:	79.51%
Best Local Similarity:	58.20%
Query Match:	56.75%
Indels:	0
Gaps:	13
Matches:	71
Conservative:	26
Mismatches:	25
Indels:	0
Gaps:	0
BASE COUNT	187 a 110 c 101 g 210 t
ORIGIN	
Alignment Scores:	
Pred. No.:	3.65e-40
Score:	374.00
Percent Similarity:	79.51%
Best Local Similarity:	58.20%
Query Match:	56.75%
Indels:	0
Gaps:	13
Matches:	71
Conservative:	26
Mismatches:	25
Indels:	0
Gaps:	0
BASE COUNT	187 a 110 c 101 g 210 t
ORIGIN	
Alignment Scores:	
Pred. No.:	3.65e-40
Score:	374.00
Percent Similarity:	79.51%
Best Local Similarity:	58.20%
Query Match:	56.75%
Indels:	0
Gaps:	13
Matches:	71
Conservative:	26
Mismatches:	25
Indels:	0
Gaps:	0
BASE COUNT	187 a 110 c 101 g 210 t
ORIGIN	
Alignment Scores:	
Pred. No.:	3.65e-40
Score:	374.00
Percent Similarity:	79.51%
Best Local Similarity:	58.20%
Query Match:	56.75%
Indels:	0
Gaps:	13
Matches:	71
Conservative:	26
Mismatches:	25
Indels:	0
Gaps:	0
BASE COUNT	187 a 110 c 101 g 210 t
ORIGIN	
Alignment Scores:	
Pred. No.:	3.65e-40
Score:	374.00
Percent Similarity:	79.51%
Best Local Similarity:	58.20%
Query Match:	56.75%
Indels:	0
Gaps:	13
Matches:	71
Conservative:	26
Mismatches:	25
Indels:	0
Gaps:	0
BASE COUNT	187 a 110 c 101 g 210 t
ORIGIN	
Alignment Scores:	
Pred. No.:	3.65e-40
Score:	374.00
Percent Similarity:	79.51%
Best Local Similarity:	58.20%
Query Match:	56.75%
Indels:	0
Gaps:	13
Matches:	71
Conservative:	26
Mismatches:	25
Indels:	0
Gaps:	0
BASE COUNT	187 a 110 c 101 g 210 t
ORIGIN	
Alignment Scores:	
Pred. No.:	3.65e-40
Score:	374.00
Percent Similarity:	79.51%
Best Local Similarity:	58.20%
Query Match:	56.75%
Indels:	0
Gaps:	13
Matches:	71
Conservative:	26
Mismatches:	25
Indels:	0
Gaps:	0
BASE COUNT	187 a 110 c 101 g 210 t
ORIGIN	
Alignment Scores:	
Pred. No.:	3.65e-40
Score:	374.00
Percent Similarity:	79.51%
Best Local Similarity:	58.20%
Query Match:	56.75%
Indels:	0
Gaps:	13
Matches:	71
Conservative:	26
Mismatches:	25
Indels:	0
Gaps:	0
BASE COUNT	187 a 110 c 101 g 210 t
ORIGIN	
Alignment Scores:	
Pred. No.:	3.65e-40
Score:	374.00
Percent Similarity:	79.51%
Best Local Similarity:	58.20%
Query Match:	56.75%
Indels:	0
Gaps:	13
Matches:	71
Conservative:	26
Mismatches:	25
Indels:	0
Gaps:	0
BASE COUNT	187 a 110 c 101 g 210 t
ORIGIN	
Alignment Scores:	
Pred. No.:	3.65e-40
Score:	374.00
Percent Similarity:	79.51%
Best Local Similarity:	58.20%
Query Match:	56.75%
Indels:	0
Gaps:	13
Matches:	71
Conservative:	26
Mismatches:	25
Indels:	0
Gaps:	0
BASE COUNT	187 a 110 c 101 g 210 t
ORIGIN	
Alignment Scores:	
Pred. No.:	3.65e-40
Score:	374.00
Percent Similarity:	79.51%
Best Local Similarity:	58.20%
Query Match:	56.75%
Indels:	0
Gaps:	13
Matches:	71
Conservative:	26
Mismatches:	25
Indels:	0
Gaps:	0
BASE COUNT	187 a 110 c 101 g 210 t
ORIGIN	
Alignment Scores:	
Pred. No.:	3.65e-40
Score:	374.00
Percent Similarity:	79.51%
Best Local Similarity:	58.20%
Query Match:	56.75%
Indels:	0
Gaps:	13
Matches:	71
Conservative:	26
Mismatches:	25
Indels:	0
Gaps:	0
BASE COUNT	187 a 110 c 101 g 210 t
ORIGIN	
Alignment Scores:	
Pred. No.:	3.65e-40
Score:	374.00
Percent Similarity:	79.51%
Best Local Similarity:	58.20%
Query Match:	56.75%
Indels:	0
Gaps:	13
Matches:	71
Conservative:	26
Mismatches:	25
Indels:	0
Gaps:	0
BASE COUNT	187 a 110 c 101 g 210 t
ORIGIN	
Alignment Scores:	
Pred. No.:	3.65e-40
Score:	374.00
Percent Similarity:	79.51%
Best Local Similarity:	58.20%
Query Match:	56.75%
Indels:	0
Gaps:	13
Matches:	71
Conservative:	26
Mismatches:	25
Indels:	0
Gaps:	0
BASE COUNT	187 a 110 c 101 g 210 t
ORIGIN	
Alignment Scores:	
Pred. No.:	3.65e-40
Score	

US-09-873-409-3 (1-131) x BJ356648 (1-608)

QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
 Db 539 ATCGATGGTCATAATAAAGAAATCTAAATGTTGTCATCTTAGACATCTCTTTGGTTA 480
 QY 22 ValSerGlnGluProValLeuPheGlyThrIleSerAsnAsnIleLysTyrGlyArg 41
 Db 479 GTTGGTCAAGAACCAACTCTATTCACTGGTACAAATGGCCGACAATATTGTTATGGTAAG 420
 QY 42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
 Db 419 CATGATGCAACTCAGAGGAGATTGAAGAGGCTTCAAAATTATCAAAATTCATTCATTTC 360
 QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
 Db 359 ATCAATTGATTTCCTCAATGGTTATATAACACAGATTGGTGGTGAGAAATACACTCAATTATCA 300
 QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
 Db 299 GGTGGTCAAAACACACGTATAGCAATTGCTCGTGCATTCATTGTAATCCAAAGATTTTA 240
 QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
 Db 239 TTATTGGATGAGTCAACCTCTGCAATTGGATGCTGATTCAACAAAATTAGTTCAAGAAGCA 180
 QY 122 LeuGlu 123
 Db 179 CTAGAA 174

Search completed: March 31, 2003, 13:50:03
 Job time : 907.167 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 30, 2003, 01:20:01 ; Search time 128.031 Seconds
(without alignments)
2304.231 Million cell updates/sec

Title: US-09-873-409-3
Perfect score: 659
Sequence: 1 MVDENDIRALNVRHYRDHIG.....SEKSAVQAALKEKTPRYSF 131

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DEV=xlh

-Q/cn2_1/USPTO_spool/US9873409/runat_27032003.115416.19216/app_query.fasta_1.7544
-DB=N Geneseq 101002 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPECL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US9873409 @CGN 1 1 4247 @runat_27032003.115416.19216 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DRV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -FgaPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 101002.*
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	619	93.9	3699	24	ABK83223	Human transporter
2	472	71.6	4233	21	AAZ90198	Rat mdrlb2 (multidr
3	472	71.6	4233	22	AAF27498	Rat mdrlb2 multidr
4	469	71.2	4279	22	AAD03488	Dog P-glycoprotein
5	469	71.2	4279	22	AAD03504	Dog P-glycoprotein
6	469	71.2	4279	22	AAD03505	Dog P-glycoprotein
7	469	71.2	4279	22	AAD03506	Dog P-glycoprotein
8	468	71.0	4189	21	AAZ49334	Murine multidrug r
9	468	71.0	4189	24	ABA94367	Mouse BCRP DNA rel
10	468	71.0	4254	14	AAQ38950	Mouse multidrug re
11	467	70.9	4254	24	ABK63517	Rat sequence diffe
12	467	70.9	4369	21	AAZ52047	Rat multidrug resi
13	467	70.9	4425	21	AAZ52048	Rat multidrug resi
14	465	70.6	4186	22	AAF86127	Cynomologous monke
15	465	70.6	4195	22	AAF86128	Cynomologous monke
16	465	70.6	4317	22	AAD03489	Dog P-glycoprotein
17	462	70.1	4788	21	AAZ49335	Murine multidrug r
18	462	70.1	4788	24	ABA94368	Mouse BCRP DNA rel
19	461	70.0	3840	24	ABL91687	Human polynucleoti
20	461	70.0	3860	21	AAZ49332	Human wild-type mu
21	461	70.0	3860	21	AAZ49333	Human G185V mutant
22	461	70.0	3860	24	ABA94365	Human BCRP DNA rel
23	461	70.0	3860	24	ABA94366	Human BCRP DNA rel
24	461	70.0	3988	21	AAZ88973	Human MDR-1 DNA..
25	461	70.0	4264	19	AAV66533	Mutated human P-gl
26	461	70.0	4264	19	AAV66534	Mutated human P-gl
27	461	70.0	4349	22	AAH57442	Human intestine ce
28	461	70.0	4378	11	AAQ04522	Multidrug Resistan
29	461	70.0	4646	15	AAQ72872	Human multidrug re
30	461	70.0	4646	21	AAZ94738	Human ATP binding
31	461	70.0	4646	24	AAD38994	Human mdri gene.
32	461	70.0	4646	24	ABL68592	Kidney cancer rela
33	461	70.0	4646	24	ABL68880	Kidney cancer rela
34	461	70.0	4669	8	AAW70752	Sequence of human
35	461	70.0	4669	14	AAQ52726	Sequence of human
36	461	70.0	4669	19	AAV32645	Human P glycoprote
37	461	70.0	4669	24	ABK52041	cDNA encoding huma
38	461	70.0	6505	17	RA113394	Hybrid vector pSF-
39	461	70.0	8630	21	AAZ24041	Retroviral M4 mdr-
40	461	70.0	8630	21	AAZ24042	Retroviral vector
41	448	68.0	3912	24	ABK63653	Rat sequence diffe
42	440	66.8	3924	21	AAZ94742	Human ATP binding
43	440	66.8	3924	21	AAZ88974	Human MDR-3 DNA.
44	440	66.8	3924	24	ABN95801	Gene #2299 used to
45	427	64.8	4776	21	AAZ94744	Human ATP binding

ALIGNMENTS

RESULT 1
ABK83223
ID ABK83223 standard; cDNA; 3699 BP.
XX
XX ABK83223;
AC
XX
XX
XX 27-AUG-2002 (first entry)
DE Human transporter and ion channel, TRICH14, Incyte ID 7472030CB1, cDNA.
XX Human; ss; gene; transporter and ion channel; TRICH; transport disorder;
KW neurological disorder; muscle disorder; immunological disorder; cancer;
KW scleroderma; systemic lupus erythematosus; allergy; leukaemia;
KW cell proliferative disorder; cervical cancer; breast cancer;
KW neurodegenerative disorder; Parkinson's disease; Alzheimer's disease;
KW myotonic dystrophy; catatonia; endocrine disorder; diabetes;
KW Grave's disease; gastrointestinal disorder; Crohn's disease;
KW renal disorder; Good pasture's syndrome; viral infection; cirrhosis;

KW bacterial infection; fungal infection; parasitic infection;
 KW protozoal infection; helminthic infection; cardiovascular disorder;
 KW atherosclerosis; hepatic disease.

OS Homo sapiens.

XX WO200240541-A2.

XX 23-MAY-2002.

XX 25-OCT-2001; 2001WO-US46055.

XX 27-OCT-2000; 2000US-243989P.

PR 03-NOV-2000; 2000US-245904P.

PR 09-NOV-2000; 2000US-247673P.

PR 17-NOV-2000; 2000US-249661P.

PR 20-NOV-2000; 2000US-252232P.

PR 01-DEC-2000; 2000US-250790P.

XX (INCY-) INCYTE GENOMICS INC.

XX Tang YT, Yue H, Nguyen DB, Hafalia AJA, Elliott VS, Lu Y;

PI Walia NK, Yao MG, Baughn MR, Gandhi AR, Ding L, Sanjanwala M;

PI Ramkumar J, Arvizu C, Gietzen KJ, Lal PG, Azimzai Y, Khan FA;

PI Thangavelu K, Thornton M, Lu DAM, Tribouley CM, Warren BA;

PI Ison CH, Das D, Raumann BE, Policky JL, Kearney L;

XX WPI: 2002-463570/49.

DR P-PSDB; ABG61544.

XX New transporters and ion channels (TRICH) polypeptides, useful for
 PT diagnosing, preventing, and treating disorders associated with an
 PT abnormal expression or activity of TRICH, e.g. immunological, muscular
 PT or renal disorders

XX Claim 5; Page 172; 178pp; English.

XX The invention relates to human transporters and ion channels (TRICH)
 CC polypeptides, a naturally occurring amino acid sequence 90 % identical to
 CC TRICH, a biologically active fragment of TRICH or an immunogenic fragment
 CC of TRICH. Also included are an isolated polynucleotide encoding TRICH,
 CC a recombinant polynucleotide comprising a promoter sequence operably
 CC linked to the TRICH polynucleotide, a cell transformed with the
 CC recombinant polynucleotide, a transgenic organism comprising the
 CC recombinant polynucleotide, an isolated antibody that binds specifically
 CC to TRICH, and screening for compounds which bind to TRICH, modulate
 CC TRICH, modulate TRICH expression or are ant/agonists of TRICH.
 CC The polypeptides are useful for diagnosing, treating, and
 CC preventing transport, neurological, muscle, immunological disorders
 CC (e.g. scleroderma, systemic lupus erythematosus, allergies), cell
 CC proliferative disorders such as cancers (e.g. leukaemia, cervical or
 CC breast cancers), neurodegenerative disorders (e.g. Parkinson's disease,
 CC Alzheimer's disease), muscular disorders (e.g. myotonic dystrophy,
 CC catatonias), endocrine disorders (e.g. diabetes, Grave's disease),
 CC gastrointestinal disorders (e.g. Crohn's disease), renal disorders
 CC (e.g. Good pasture's syndrome), viral, bacterial, fungal, parasitic,
 CC protozoal and helminthic infections, cardiovascular disorders (e.g.
 CC atherosclerosis), or hepatic diseases (e.g. cirrhosis) and many
 CC other diseases and disorders detailed in the specification. They can also
 CC be used in assessing the effects of exogenous compounds on the
 CC expression of nucleic acid and amino acid sequences of transporters and
 CC ion channels. TRICH or its fragments may also be used in screening for
 CC compounds that specifically bind to and modulate the activity of TRICH.
 CC The polynucleotides can be used to create knock-in humanised animals or
 CC transgenic animals to model human disease. The present sequence
 CC encodes a TRICH protein.

XX Sequence 3699 BP; 1116 A; 707 C; 860 G; 1016 T; 0 other;

Alignment Scores:

Pred. No.: 3.68e-72 Length: 3699
 Score: 619.00 Matches: 124
 Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 93.93% Indels: 0
 DB: 24 Gaps: 0
 US-09-873-409-3 (1-131) x ABK83223 (1-3699)
 QY 1 MetValAspGluAenAspIleArgAlaLeuAenValArgHisTyrArgAspHisIleGly 20
 DB 1339 ATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGGCATTATCGAGACCATTATTGGA 1398
 QY 21 ValValSerGlnGluProValLeuPheGlyThrIleSerAsnAsnIleLysTyrGly 40
 DB 1399 GTGGTTAGTCAAGAGCGCTGTTTTGTTCCGGACCACCATCATCAATATCAAGTATGGA 1458
 QY 41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
 DB 1459 CGAGATGATGTGACTGATGAAGAGTGGAGAGCAGCAGGAGCAAGGAGCAATGGTATGAT 1518
 QY 61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80
 DB 1519 TTTATCATCGAGTTTCTTAATAAATTTAATACATTGGTAGGGGAAAAAGAGCTCAATG 1578
 QY 81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100
 DB 1579 AGTGGAGGGCAGAAACAGAGGATCGCAATTGCTCGCTTAGTTCGAAACCCCAAGATT 1638
 QY 101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120
 DB 1639 CTGATTTTATGATGAGGCTAGCTCTGCCCTGGATTGAGAAAGCAAGTCAGCTGTTCAAGCT 1698
 QY 121 AlaLeuGluLys 124
 DB 1699 GCACCTGGAGAAG 1710

RESULT 2

AAZ90198
 ID AAZ90198 standard; DNA; 4233 BP.

XX AC AAZ90198;

XX DT 19-MAY-2000 (first entry)

XX DE Rat mdrlb2 (multispecific drug transporter) nucleotide sequence.

XX KW Rat; mdrlb2; multispecific drug transporter; cancer; inflammation;
 KW cardiovascular disease; central nervous system disorder;
 KW autoimmune disease; kidney disease; drug formulation; ss.

XX OS Rattus sp.

XX PN US6025160-A.

XX PD 15-FEB-2000.

XX PF 22-JUL-1998; 98US-0120513.

XX PR 22-JUL-1998; 98US-0120513.

XX SM (SMIK) SMITHKLINE BEECHAM CORP.

XX PI Brun KA, Ellens H, Yue L, Chenery RJ, Feild JA;

XX WPI: 2000-181810/16.

XX DR P-PSDB; AAY78879.

XX Isolated polynucleotide encoding a rat mdrlb2 drug transporter
 PT polypeptide, useful for treatment of e.g. cancer, autoimmune disease,
 PT central nervous system disorders

XX PS Claim 7; Column 13-17; 17pp; English.

XX CC This sequence represents a polynucleotide encoding a 1275 amino acid
 CC rat mdrlb2 multispecific drug transporter polypeptide. The mdrlb2

CC nucleotide sequence has cytostatic, antiinflammatory, cardiant,
 CC neuroprotective, immunosuppressive and nephrotropic activity.
 CC Understanding the functioning of the mdrlb2 polynucleotide and protein in
 CC transgenic animal models is useful for treating and preventing diseases
 CC such as cancer, inflammation, cardiovascular disease, central nervous
 CC system disorders, autoimmune disease, and kidney disease. The use of the
 CC protein in cell based, membrane based, or binding assays may enhance drug
 CC formulation, selection of formulation excipients and compound design.
 XX

SQ Sequence 4233 BP; 1210 A; 920 C; 1048 G; 1055 T; 0 other;

Alignment Scores: 1.96e-52 Length: 4233
 Pred. No.: 472.00 Matches: 91
 Score: 91.06% Conservative: 21
 Best Local Similarity: 73.98% Mismatches: 11
 Query Match: 71.62% Indels: 0
 DB: 21 Gaps: 0

US-09-873-409-3 (1-131) x AA290198 (1-4233)

QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
 DB 1376 ATCGACGGACAGACATCAGACCATCAATGTGAGGTATCTGCGGAATCATTTGGGGTG 1435
 QY 22 ValSerGlnGluProValLeuPheGlyThrIleSerAsnAsnIleLysTyrGlyArg 41
 DB 1436 GTGAGTCAGAACCCGCTGCTTTTGCACCATGTCGCGAAACATTCGCTATGCGCGA 1495
 QY 42 AspAspValThrAspGluMetGluArgAlaAArgGluAlaAsnAlaTyrAspPhe 61
 DB 1496 GAAACGTCACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1555
 QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGlyAlaGlnMetSer 81
 DB 1556 ATCATGAATCGCCACAAATTTAAACCCCTGTTGTTGAGAGAGGGCGCAGCTGAGT 1615
 QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaAArgAlaLeuValArgAsnProLysIleLeu 101
 DB 1616 GGGGGACAGAAACAGAGGATCGCATTTGCCGGCCCTGTCGCAACCCCAAGATCCTT 1675
 QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaA 121
 DB 1676 TTGTTGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGGCGTGTTCAGCGCGCT 1735
 QY 122 LeuGlyLys 124
 DB 1736 CTGGATAAG 1744

RESULT 3
 AAF27498
 ID AAF27498 standard; CDNA; 4233 BP.

XX AAF27498;
 XX
 XX 25-APR-2001 (first entry)
 XX
 XX Rat mdrlb2 multidrug resistance CDNA, SEQ ID NO:1.
 XX
 XX Rat; multidrug resistance; mdrlb2; multi-drug transporter family;
 KW drug screening; pharmacokinetic analysis; oral absorption;
 KW formulation design; bioavailability; transgenic animal; knockout animal;
 KW inflammation; cardiovascular disease; central nervous system disorder;
 KW cancer; autoimmune disorder; kidney disease; ss.
 XX
 XX Rattus sp.
 XX
 XX US6169166-B1.
 XX
 XX 02-JAN-2001.
 XX
 XX 29-NOV-1999; 99US-0450105.
 XX
 XX

PR 22-JUL-1998; 98US-0120513.
 XX (SMIK) SMITHKLINE BEECHAM CORP.
 PA (SMIK) SMITHKLINE BEECHAM PLC.
 XX
 PI Chenery RJ, Ellens H, Brun KA, Yue L, Peild JA;
 XX WPI; 2001-158280/16.
 DR P-PSDB; AAB60409.
 XX
 PT New rat multidrug resistance protein (mdrlb2), useful for screening or
 PT identifying compounds that are (ant)agonists/inhibitors of the
 PT ratmdrlb2, as well as compounds with optimal development
 PT characteristics -
 XX
 PS Disclosure; Column 13-14; 13pp; English.
 XX
 CC The invention relates to a novel rat multidrug resistance (mdr)
 CC protein, mdrlb2 (AAB60409). Multi-specific drug transporter family
 CC proteins are present in cell which have a barrier function, such as
 CC intestinal epithelial cells, brain microvessel endothelial cells,
 CC kidney epithelial cells, and liver hepatocytes, and are also expressed
 CC by certain cancer cells. The rat mdrlb2 protein is useful for
 CC screening or identifying compounds that are agonists or antagonists
 CC of mdrlb2 activity. It may also be used to establish assays to
 CC predict oral absorption and pharmacokinetics of drugs in humans, and
 CC thus enhance the design of formulations through the identification of
 CC compounds with optimal development characteristics (i.e., high oral
 CC bioavailability, UID (once a day) dosing, reduced drug interactions,
 CC reduced variability, and reduced food effects), specifically to
 CC avoid interactions with human mdr-1. Transgenic and knockout animals
 CC created using DNA encoding the rat mdrlb2 may be used to gain an
 CC insight into treating and preventing human diseases such as cancer,
 CC inflammation, cardiovascular disease, central nervous system disorders,
 CC autoimmune disorders and kidney disease. The present sequence represents
 CC cDNA encoding rat mdrlb2.
 XX
 SQ Sequence 4233 BP; 1210 A; 920 C; 1048 G; 1055 T; 0 other;

Alignment Scores: 1.96e-52 Length: 4233
 Pred. No.: 472.00 Matches: 91
 Score: 91.06% Conservative: 21
 Best Local Similarity: 73.98% Mismatches: 11
 Query Match: 71.62% Indels: 0
 DB: 22 Gaps: 0
 US-09-873-409-3 (1-131) x AAF27498 (1-4233)
 QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
 DB 1376 ATCGACGGACAGACATCAGACCATCAATGTGAGGTATCTGCGGAATCATTTGGGGTG 1435
 QY 22 ValSerGlnGluProValLeuPheGlyThrIleSerAsnAsnIleLysTyrGlyArg 41
 DB 1436 GTGAGTCAGAACCCGCTGCTTTTGCACCATGTCGCGAAACATTCGCTATGCGCGA 1495
 QY 42 AspAspValThrAspGluMetGluArgAlaAArgGluAlaAsnAlaTyrAspPhe 61
 DB 1496 GAAACGTCACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1555
 QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGlyAlaGlnMetSer 81
 DB 1556 ATCATGAATCGCCACAAATTTAAACCCCTGTTGTTGAGAGAGGGCGCAGCTGAGT 1615
 QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaAArgAlaLeuValArgAsnProLysIleLeu 101
 DB 1616 GGGGGACAGAAACAGAGGATCGCATTTGCCGGCCCTGTCGCAACCCCAAGATCCTT 1675
 QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaA 121
 DB 1676 TTGTTGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGGCGTGTTCAGCGCGCT 1735


```

QY 122 LeuGluLys 124
Db 1736 CTGGATAAG 1744

RESULT 4
AAD03488
ID AAD03488 standard; cDNA; 4279 BP.
AC AAD03488;
XX 13-JUN-2001 (first entry)
DE Dog P-glycoprotein (PGP) cDNA #1.
XX
KW Dog; P-glycoprotein; PGP; multidrug transporter; MDR1;
KW drug bioavailability; transgenic animal; genetic model; ss.
XX
OS Canis familiaris.
XX
FH Key Location/Qualifiers
FT CDS 17..3862
FT /tag= a
FT /product= "Dog P-glycoprotein (PGP) #1"
XX
WO200123540-A2.
XX
PD 05-APR-2001.
PF 28-SEP-2000; 2000WO-US26767.
XX
PR 28-SEP-1999; 99US-0156510.
XX
PA (GENT-) GENTEST CORP.
XX
PI Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;
XX
DR WPI; 2001-235373/24.
XX
DR P-PSDB; AAE00303.
XX
PT New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
PT for determining the bioavailability of drugs and for screening for dog
PT PGP inhibitors -
XX
PS Claim 3; Page 58-63; 111pp; English.
XX
CC The invention relates to dog P-glycoprotein (PGP) also referred
CC as multidrug transporter (MDR1) and nucleic acids encoding them.
CC The invention also includes fragments and biologically functional
CC variants of dog P-glycoprotein. PGP and their nucleic acids are
CC useful for determining the bioavailability of drugs and for
CC screening PGP inhibitors. They are useful for the diagnosis and
CC treatment of conditions characterised by PGP activity, by
CC reducing or increasing PGP activity in a cell. PGP nucleic acids
CC are used as oligonucleotide probes. Complements of PGP nucleic
CC acids are useful as antisense oligonucleotides, to induce a PGP
CC 'knockout' phenotype. They are used to prepare a non-human
CC transgenic animal, which are valuable as genetic models for
CC human diseases.
CC The present sequence is dog P-glycoprotein (PGP) cDNA. This
CC sequence is also referred as Genotype C cDNA. The
CC PGP enzyme functions as an efflux pump exporting small molecules
CC across the cell membrane. This enzyme is a member of the ABC
CC transporter family.
XX
SQ Sequence 4279 BP; 1294 A; 834 C; 1008 G; 1143 T; 0 other;

Alignment Scores:
Pred. No.: 5e-52 Length: 4279
Score: 469.00 Matches: 90
Percent Similarity: 89.43% Conservative: 20
Best Local Similarity: 73.17% Mismatches: 13
Query Match: 71.17% Indels: 0
DB: 22 Gaps: 0

```

```

US-09-873-409-3 (1-131) x AAD03488 (1-4279)
QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValAtcGHisTyrArgAspHisIleGlyVal 21
Db 1379 ATTGATGGACAGACATTAGGACCATAAATGTAAGCACTCTTCGGGAATATTACTGTGTG 1438
QY 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
Db 1439 GTGAGTCAGAGGCTGTGTTGTCACACAGATAGTGAACACATTGCTATGCGCCG 1498
QY 42 AspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
Db 1499 GAAATGTCACCATGATGAGATTGAGAAAGCTGTTAAGAGAGCCCAATGCTATGATTTT 1558
QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
Db 1559 ATCATGAAACTACCTAATAATTTGACACTCTGTTGGAGAGAGAGGGCCCACTGAGT 1618
QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
Db 1619 GGTGGACAGAAACAGAGAAATCGCCATTGCTCGGGCCCTGTTGCAACCCCAAGATTTCT 1678
QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
Db 1679 CTGCTGGATGAGGCAACGTCAGCTCTGGACACTGGAAGAGTGAAGCAGTGGTTCAGTGGCC 1738
QY 122 LeuGluLys 124
Db 1739 CTGGATAAG 1744

RESULT 5
AAD03504
ID AAD03504 standard; cDNA; 4279 BP.
XX
AC AAD03504;
XX
DT 13-JUN-2001 (first entry)
XX
DE Dog P-glycoprotein (PGP) allelic variant (Genotype A) cDNA.
XX
KW Dog; P-glycoprotein allelic variant; multidrug transporter; MDR1;
KW drug bioavailability; transgenic animal; genetic model; ss.
XX
OS Canis familiaris.
XX
FH Key Location/Qualifiers
FT CDS 17..3862
FT /tag= a
FT /product= "Dog P-glycoprotein (PGP) allelic variant
FT allele (Genotype A) protein"
FT replace (607, C)
FT /*tag= b
XX
WO200123540-A2.
XX
PD 05-APR-2001.
PF 28-SEP-2000; 2000WO-US26767.
XX
PR 28-SEP-1999; 99US-0156510.
XX
PA (GENT-) GENTEST CORP.
XX
PI Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;
XX
DR WPI; 2001-235373/24.
XX
DR P-PSDB; AAE00308.
XX
PT New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
PT for determining the bioavailability of drugs and for screening for dog
PT PGP inhibitors -
XX
XX

```


Claim 9; Page 85-90; 111pp; English.

PS The invention relates to dog P-glycoprotein (PGP) also referred
 XX as multidrug transporter (MDR1) and nucleic acids encoding them.
 CC The invention also includes fragments and biologically functional
 CC variants of dog P-glycoprotein. PGP and their nucleic acids are
 CC useful for determining the bioavailability of drugs and for
 CC screening PGP inhibitors. They are useful for the diagnosis and
 CC treatment of conditions characterised by PGP activity, by
 CC reducing or increasing PGP activity in a cell. PGP nucleic acids
 CC are used as oligonucleotide probes. Complements of PGP nucleic
 CC acids are useful as antisense oligonucleotides, to induce a PGP
 CC 'knockout' phenotype. They are used to prepare a non-human
 CC transgenic animal, which are valuable as genetic models for
 CC human diseases.
 CC The present sequence is dog P-glycoprotein (PGP) allelic variant
 CC (Genotype A) cDNA. The PGP enzyme functions as an efflux pump
 CC exporting small molecules across the cell membrane. This enzyme
 CC is a member of the ABC transporter family.
 XX

SQ Sequence 4279 BP; 1295 A; 833 C; 1008 G; 1143 T; 0 other;

Alignment Scores:
 Pred. No.: 5e-52 Length: 4279
 Score: 469.00 Matches: 90
 Percent Similarity: 89.43% Conservative: 20
 Best Local Similarity: 73.17% Mismatches: 13
 Query Match: 71.17% Indels: 0
 DB: 22 Gaps: 0

US-09-873-409-3 (1-131) x AAD03504 (1-4279)

QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
 Db 1379 ATTGATGGACAGGACATAGGACATTAATGTAAAGCACTTCGGGAATATTCTGGTGTG 1438
 QY 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
 Db 1439 GTGAGTCAGGAGCGCTGTGTGTTTGGCCACCATGAGTCTGAAACATTCGCTATGCGCGC 1498
 QY 42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
 Db 1499 GAAATATGTCACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1558
 QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
 Db 1559 ATCATGAATCTACTATAATTTTGACACTCTGTTGGAGAGAGAGGGGCCAGCTGAGT 1618
 QY 82 GlyGlyGlnLysGlnArgIleAlaAlaArgAlaLeuValArgAsnProLysIleLeu 101
 Db 1619 GGTGGACAGAAACAGAGAAATCGCATTTGTCGGGCCCTGCTGCAACCCCAAGATTCTT 1678
 QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
 Db 1679 CTGCTGGATGAGGACAGCTCAGCTCTGACACTCTGCTGAGAGAGAGGGGCCAGCTGAGT 1738
 QY 122 LeuGluLys 124
 Db 1739 CTGGATAAG 1747
 RESULT 6
 AAD03505
 ID AAD03505 standard; cDNA; 4279 BP.
 XX
 AC AAD03505;
 XX
 DT 13-JUN-2001 (first entry)
 XX
 DE Dog P-glycoprotein (PGP) allelic variant (Genotype B) cDNA.
 XX
 KW Dog; P-glycoprotein allelic variant; PGP; multidrug transporter;
 KW MDR1; drug bioavailability; transgenic animal; genetic model; ss.
 XX

Canis familiaris.

Key Location/Qualifiers
 CDS 17..3862
 /*tag= a
 /product= "Dog P-glycoprotein (PGP) allelic variant
 (Genotype B) protein"
 allele replace (91, T)
 /*tag= b
 allele replace (607, C)
 /*tag= c

WO200123540-A2.

05-APR-2001.

28-SEP-2000; 2000WO-US26767.

28-SEP-1999; 99US-0156510.

(GENT-) GENTEST CORP.

Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;

WPI; 2001-235373/24.

P-PSDB; AAE00309.

New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
 for determining the bioavailability of drugs and for screening for dog
 PGP inhibitors -

Claim 9; Page 93-99; 111pp; English.

The invention relates to dog P-glycoprotein (PGP) also referred
 as multidrug transporter (MDR1) and nucleic acids encoding them.
 The invention also includes fragments and biologically functional
 variants of dog P-glycoprotein. PGP and their nucleic acids are
 useful for determining the bioavailability of drugs and for
 screening PGP inhibitors. They are useful for the diagnosis and
 treatment of conditions characterised by PGP activity, by
 reducing or increasing PGP activity in a cell. PGP nucleic acids
 are used as oligonucleotide probes. Complements of PGP nucleic
 acids are useful as antisense oligonucleotides, to induce a PGP
 'knockout' phenotype. They are used to prepare a non-human
 transgenic animal, which are valuable as genetic models for
 human diseases.
 The present sequence is dog P-glycoprotein (PGP) allelic variant
 (Genotype B) cDNA. The PGP enzyme functions as an efflux pump
 exporting small molecules across the cell membrane. This enzyme
 is a member of the ABC transporter family.

Sequence 4279 BP; 1296 A; 833 C; 1008 G; 1142 T; 0 other;

Alignment Scores:

Pred. No.: 5e-52 Length: 4279
 Score: 469.00 Matches: 90
 Percent Similarity: 89.43% Conservative: 20
 Best Local Similarity: 73.17% Mismatches: 13
 Query Match: 71.17% Indels: 0
 DB: 22 Gaps: 0

US-09-873-409-3 (1-131) x AAD03505 (1-4279)

QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
 Db 1379 ATTGATGGACAGGACATAGGACATTAATGTAAAGCACTTCGGGAATATTCTGGTGTG 1438
 QY 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
 Db 1439 GTGAGTCAGGAGCGCTGTGTGTTTGGCCACCATGAGTCTGAAACATTCGCTATGCGCGC 1498
 QY 42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
 Db 1499 GAAATATGTCACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1558

Db 1499 GAAATGTCACCATGGATGAGATTGAGAAAGCTGTTAAGGAAGCCAAATGCTATGATTTT 1558
 QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
 Db 1559 ATCATGAACTACCTATATAATTTGACACTCTGTTGGAGAGAGAGGGCCAGCTGAGT 1618
 QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
 Db 1619 GGTGGACAGAAACAGAGAAATCGCATTTGCTCGGGCCCTGGTTGCAACCCCAAGATTCTT 1678
 QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
 Db 1679 CTGCTGGATGAGGCAACGTGAGTCTCTGGACACTGAAAGTGAAGCAGTGGTTTCAGGTGCC 1738
 QY 122 LeuGluLys 124
 Db 1739 CTGGATAAG 1747
 RESULT 7
 AAD03506
 ID AAD03506 standard; cDNA; 4279 BP.
 XX
 AC AAD03506;
 XX
 DT 13-JUN-2001 (first entry)
 DE Dog P-glycoprotein (PGP) allelic variant (Genotype D) cDNA.
 XX
 KW Dog; P-glycoprotein allelic variant; PGP; multidrug transporter;
 KW MDRI; drug bioavailability; transgenic animal; genetic model; ss.
 XX
 OS Canis familiaris.
 XX
 FH Key Location/Qualifiers
 FT CDS 17..3862
 FT /*tag= a
 FT /product= "Dog P-glycoprotein (PGP) allelic variant
 (Genotype D) protein"
 FT allele replace (91, T)
 FT /*tag= b
 FT allele replace (607, C)
 FT /*tag= c
 FT allele replace (1001, T)
 FT /*tag= c
 FT allele replace (3458, A)
 FT /*tag= c
 XX
 PN WO200123540-A2.
 XX
 PD 05-APR-2001.
 XX
 PP 28-SEP-2000; 2000WO-US26767.
 XX
 PR 28-SEP-1999; 99US-0156510.
 XX
 PA (GENT-) GENTEST CORP.
 XX
 PI Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;
 XX
 DR WPI; 2001-235373/24.
 DR P-PSDB; AAE00310.
 XX
 PT New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
 for determining the bioavailability of drugs and for screening for dog
 PGP inhibitors -
 XX
 PS Claim 9; Page 102-107; Illipp; English.
 XX
 CC The invention relates to dog P-glycoprotein (PGP) also referred
 as multidrug transporter (MDRI) and nucleic acids encoding them.
 CC The invention also includes fragments and biologically functional
 variants of dog P-glycoprotein. PGP and their nucleic acids are
 useful for determining the bioavailability of drugs and for

CC screening PGP inhibitors. They are useful for the diagnosis and
 CC treatment of conditions characterised by PGP activity, by
 CC reducing or increasing PGP activity in a cell. PGP nucleic acids
 CC are used as oligonucleotide probes. Complements of PGP nucleic
 CC acids are useful as antisense oligonucleotides, to induce a PGP
 CC 'knockout' phenotype. They are used to prepare a non-human
 CC transgenic animal, which are valuable as genetic models for
 CC human diseases.
 CC The present sequence is dog P-glycoprotein (PGP) allelic variant
 CC (Genotype D) cDNA. The PGP enzyme functions as an efflux pump
 CC exporting small molecules across the cell membrane. This enzyme
 CC is a member of the ABC transporter family.
 XX
 SQ Sequence 4279 BP; 1296 A; 833 C; 1009 G; 1141 T; 0 other;
 Alignment Scores:
 Pred. No.: 5e-52 Length: 4279
 Score: 469.00 Matches: 90
 Percent Similarity: 89.43% Conservative: 20
 Best Local Similarity: 73.17% Mismatches: 13
 Query Match: 71.17% Indels: 0
 DB: 22 Gaps: 0
 US-09-873-409-3 (1-131) x AAD03506 (1-4279)
 QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
 Db 1379 ATTGATGGACAGACATTAGGACCATAAATGTAAGGCATCTTCGGGAAATTAATCTGGTGTG 1438
 QY 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
 Db 1439 GTGAGTCAGAGCCTGTGTGTTTGGCCACCATGATAGCTGAAACCATTCGCTATGCGCCG 1498
 QY 42 AspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
 Db 1499 GAAATGTCACCATGGATGAGATTGAGAAAGCTGTTAAGGAAGCCAAATGCTATGATTTT 1558
 QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
 Db 1559 ATCATGAACTACCTATATAATTTGACACTCTGTTGGAGAGAGAGGGCCAGCTGAGT 1618
 QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
 Db 1619 GGTGGACAGAAACAGAGAAATCGCATTTGCTCGGGCCCTGGTTGCAACCCCAAGATTCTT 1678
 QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
 Db 1679 CTGCTGGATGAGGCAACGTGAGTCTCTGGACACTGAAAGTGAAGCAGTGGTTTCAGGTGCC 1738
 QY 122 LeuGluLys 124
 Db 1739 CTGGATAAG 1747
 RESULT 8
 AAZ49334
 ID AAZ49334 standard; cDNA; 4189 BP.
 XX
 AC AAZ49334;
 XX
 DT 14-MAR-2000 (first entry)
 DE Murine multidrug resistance-1 (MDR-1) cDNA.
 XX
 KW Multidrug resistance; MDR-1; P-glycoprotein;
 KW transmembrane efflux pump; haematopoietic stem cell; transduction;
 KW bone marrow transplantation; chemotherapy; radiation therapy; cancer;
 KW gene therapy; gene replacement; genetic defect; thalassaemia;
 KW Gauchier's disease; sickle cell anaemia; leukaemia; ex vivo expansion;
 KW cytokine; ds.
 XX
 OS Mus sp.
 XX
 FH Key Location/Qualifiers

FT CDS 1..3831
 FT /*tag= a
 FT /product= "Murine MDR-1 protein"
 PN W0961589-A2.
 XX
 XX
 PD 02-DEC-1999.
 XX
 XX 27-MAY-1999; 99WO-US11825.
 PF 28-MAY-1998; 98US-0086988.
 XX
 PR
 XX (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
 PA
 XX Sorrentino B, Bunting K;
 PI WPI: 2000-072615/06.
 DR P-PSDB; AAY58188.
 DR
 XX
 XX Ex vivo expansion of hematopoietic stem cells transduced with a
 PT sequence encoding human multidrug resistance-1, used for bone marrow
 PT transplantation -
 PS Disclosure; Page 90-93; 113pp; English.
 XX
 CC This sequence represents cDNA encoding murine multidrug
 CC resistance protein MDR-1. MDR-1 is a transmembrane
 CC efflux pump, responsible for the export of drugs from cells,
 CC particularly cancer cells. The invention relates to transducing
 CC hematopoietic stem cells with nucleic acid encoding an MDR protein
 CC and culturing the modified cells. The modified hematopoietic stem
 CC cells are useful in bone marrow transplantation (to reconstitute
 CC hematopoietic systems in patients who have undergone chemotherapy or
 CC radiation therapy) and in ex vivo gene therapy of genetic defects in
 CC cells derived from hematopoietic stem cells, e.g. thalassemia,
 CC Gauchier's disease, sickle cell anaemia or leukaemia. The modified
 CC cells can also be used to identify factors involved in regulating
 CC proliferation and differentiation in hematopoietic stem cells.
 CC Hematopoietic stem cells that express MDR-1 will be protected against
 CC chemotherapeutic agents, so can be engrafted while the patient is
 CC undergoing chemotherapy. Expansion of (rare) hematopoietic stem cells
 CC provides sufficient cells to permit standard biochemical analysis.
 CC Overexpression of MDR-1 allows cytokine-driven expansion of
 CC hematopoietic stem cells by at least 10-fold compared with a maximum
 CC of 4-fold in known procedures.
 XX
 XX Sequence 4189 BP; 1204 A; 875 C; 1028 G; 1082 T; 0 other;

Alignment Scores:

Pred. No.: 6.6e-52 Length: 4189
 Score: 468.00 Matches: 90
 Percent Similarity: 91.06% Conservative: 22
 Best Local Similarity: 73.17% Mismatches: 11
 Query Match: 71.02% Indels: 0
 DB: 21 Gaps: 0

US-09-873-409-3 (1-131) x AAZ49334 (1-4189)

QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
 DB 1354 ATCCGCGGCAAGACATCAGAACCAATCAATGTGAGGTATCTGAGGGGAGATCAATGGGTG 1413
 QY 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAniLeIysTyrGlyArg 41
 DB 1414 GTGAGTCAGAACTGTGCTGTTTGGCCACACCATCCGAGAACATTCCTATGGCCGA 1473
 QY 42 AspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAniLaTyrAspPhe 61
 DB 1474 GAAGATGTCAACATGGATGAGATTGAAAGCTCTCAAGGAGCAATGCTATGACTTC 1533
 QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
 DB 1534 ATCATGAAACTGCCCCCAATTTGACACCTGTGTTGGTGGAGAGGGGCGCAGTGTGAGT 1593

QY 82 GlyGlyGlnIysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
 DB 1594 GGGGGACAGAAACAGAGATCCCATTTGCCGGGCCCTGTGTCGCAATCCCAAGATCCTT 1653
 QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
 DB 1654 TTGTTGGACGAGGCCACCTCAGCCCTGGATACAGAAAGTGAAGCTGTGTGACGGCCGCA 1713
 QY 122 LeuGluLys 124
 DB 1714 CTGGATAAG 1722
 RESULT 9
 ABA94367
 ID ABA94367 standard; DNA; 4189 BP.
 XX
 AC ABA94367;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Mouse BCRP DNA related seq Id No. 5.
 XX
 KW Stem cell; ATP transport protein; ATP-binding cassette; antiparkinsonian;
 KW hepatotropic; neurodegenerative; cytosolic; antianemic; muscular; BCRP;
 KW cardiant; gene therapy; ds.
 OS Mus musculus.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..3831
 FT /*tag= a
 XX
 PN W0200192877-A2.
 PD 06-DEC-2001.
 XX
 PF 30-MAY-2001; 2001WO-US17459.
 XX
 PR 31-MAY-2000; 2000US-0584586.
 PR 29-MAY-2001; 2001US-0866866.
 XX
 XX (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
 PA
 PI Sorrentino B, Schuetz J;
 DR WPI: 2002-114368/15.
 DR P-PSDB; ABB07268.
 XX
 PT Identifying a stem cell, for treating e.g., muscular dystrophy,
 PT myocardial infarction, Parkinson's disease, or neurodegenerative
 PT disorders, comprises detecting the expression of an ATP transport
 PT protein (BCRP) by a cell -
 XX
 PS Disclosure; Page 64-66; 87pp; English.
 XX
 CC The invention provides a method of identifying and/or isolating a stem
 CC cell that involves detecting the expression of an ATP transport protein
 CC containing a conserved ATP-binding cassette (BCRP) by a cell in a sample
 CC comprising stem cells. The isolated stem cells may be used in the
 CC treatment of diseases such as muscular dystrophy, degenerative liver
 CC disorders, myocardial infarction, Parkinson's disease, degenerative
 CC disorders of the brain, and for tissue regeneration or replacement.
 CC Hematopoietic cells can be used in bone marrow transplants (e.g., for
 CC treatment of leukemia) and for ex vivo gene therapy for treating blood
 CC diseases such as sickle cell anemia and thalassemia. The stem cells can
 CC also be used as cell targets in gene therapy protocols. The present
 CC sequence represents a sequence related to the BCRP for which no relevant
 CC information has been provided in the specification.
 XX
 SQ Sequence 4189 BP; 1204 A; 875 C; 1028 G; 1082 T; 0 other;

Alignment Scores:


```

Pred. No.: 6.6e-52 Length: 4189
Score: 468.00 Matches: 90
Percent Similarity: 91.06% Conservative: 22
Best Local Similarity: 73.17% Mismatches: 11
Query Match: 71.02% Indels: 0
DB: 24 Gaps: 0

US-09-873-409-3 (1-131) x ABA94367 (1-4189)

QY 2 ValAspGluAsnAspIleAArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
DB 1354 ATCCAGCGGACAAAGACATCAATGTGAGGTATCTGAGGGAGATCATTTGGTGTG 1413
QY 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
DB 1414 GTGAGTCAGGAACTGTGCTGTTTGCACACAGATGCCGAGAACATTCGCTATGGCCGA 1473
QY 42 AspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
DB 1474 GAAGATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCAATGCTATGACTTC 1533
QY 62 IleMetGluPheProAsnLysPheAsnThrIleuValGlyGluLysGlyAlaGlnMetSer 81
DB 1534 ATCATGAAACTGCCCCCAATTTGACACCTGTTGTTGAGAGAGGGGGCGAGCTGAGT 1593
QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
DB 1594 GGGGACAGAAACAGAGATCGCATTTGCCGGGCTGTTGCCAATCCAGATCCTT 1653
QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
DB 1654 TTGTTGACGAGGCGCCTCAGCCCTCGATACAGAAAGTGAAGCTGTGTCAGGCGCGCA 1713
QY 122 LeuGluLys 124
DB 1714 CTGGATAAG 1722

RESULT 10
AAQ38950 standard; DNA; 4313 BP.
XX AAQ38950;
AC AAQ38950;
XX
DT 28-JUL-1993 (first entry)
DE Mouse multidrug resistance sequence.
XX
KW mdr gene; Lambda DR11 clone; ss.
OS Mus musculus.
FH Key Location/Qualifiers
FT CDS 110..3940
FT FT /*tag= a
FT FT /phenotype= multidrug_resistance
XX
XX
XX US5198344-A.
XX
XX 30-MAR-1993.
XX
XX 15-JUL-1986; 86US-0885951.
XX
XX 15-JUL-1986; 86US-0885951.
XX
XX 06-FEB-1991; 91US-0652311.
XX
XX (MASI ) MASSACHUSETTS INST TECHNOLOGY.
XX
XX Croop JM, Gros P, Housman DE;
XX
XX WPI; 1993-126077/15.
XX
XX P-PSDB; AAR35199.
XX
XX DNA sequence which confers multi-drug resistance on sensitive

```

```

PT mammalian cells - used to preserve bone marrow cells during
PT chemotherapy to prevent infection
XX
PS Claim 1; Fig 8; 22pp; English.
XX
CC A cDNA library was constructed from mRNA which had been isolated
CC from a drug-sensitive mouse cell-line. Two mouse cDNA molecules
CC complementary to the mRNA species encoded by 2 related but distinct
CC mdr genes were isolated and cloned. One of the cDNA clones (lambda
CC DR11) is a full-length cDNA clone for one member of the mdr gene
CC family. When incorporated into prokaryotic expression vector pDREX4
CC (which allows high levels of transcription of the cDNA when
CC introduced into mammalian cells), the clone was shown to confer the
CC multidrug resistance phenotype upon transfection into drug-sensitive
CC mammalian cells.
XX
SQ Sequence 4313 BP; 1241 A; 904 C; 1058 G; 1110 T; 0 other;

Alignment Scores:
Pred. No.: 6.87e-52 Length: 4313
Score: 468.00 Matches: 90
Percent Similarity: 91.06% Conservative: 22
Best Local Similarity: 73.17% Mismatches: 11
Query Match: 71.02% Indels: 0
DB: 14 Gaps: 0

US-09-873-409-3 (1-131) x AAQ38950 (1-4313)

QY 2 ValAspGluAsnAspIleAArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
DB 1463 ATCCAGCGGACAAAGACATCAATGTGAGGTATCTGAGGGAGATCATTTGGTGTG 1522
QY 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
DB 1523 GTGAGTCAGGAACTGTGCTGTTTGCACACAGATGCCGAGAACATTCGCTATGGCCGA 1582
QY 42 AspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
DB 1583 GAAGATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCAATGCTATGACTTC 1642
QY 62 IleMetGluPheProAsnLysPheAsnThrIleuValGlyGluLysGlyAlaGlnMetSer 81
DB 1643 ATCATGAAACTGCCCCCAATTTGACACCTGTTGTTGAGAGAGGGGGCGAGCTGAGT 1702
QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
DB 1703 GGGGACAGAAACAGAGATCGCATTTGCCGGGCTGTTGCCAATCCAGATCCTT 1762
QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
DB 1763 TTGTTGACGAGGCGCCTCAGCCCTCGATACAGAAAGTGAAGCTGTGTCAGGCGCGCA 1822
QY 122 LeuGluLys 124
DB 1823 CTGGATAAG 1831

RESULT 11
ABK63517
ID ABK63517 standard; cDNA; 4254 BP.
XX
XX ABK63517;
AC
XX
XX 18-JUN-2002 (first entry)
XX
XX Rat sequence differentially expressed in response to a hepatotoxin #1424.
XX
XX Rat; ss; hepatotoxin; expressed sequence tag; EST; drug screening;
XX
XX differential expression; centrilobular necrosis; steatosis.
XX
XX Rattus norvegicus.
XX
XX
XX WO200210453-A2.
XX
XX

```


PD 07-FEB-2002.

XX 30-JUL-2001; 2001WO-US23872.

XX 31-JUL-2000; 2000US-222040P.

PR 02-NOV-2000; 2000US-244880P.

PR 11-MAY-2001; 2001US-290029P.

PR 15-MAY-2001; 2001US-290645P.

PR 22-MAY-2001; 2001US-292336P.

PR 06-JUN-2001; 2001US-295798P.

PR 13-JUN-2001; 2001US-297457P.

PR 19-JUN-2001; 2001US-298884P.

PR 09-JUL-2001; 2001US-303459P.

XX (GENE-) GENE LOGIC INC.

XX Mendrick D, Porter MW, Johnson KR, Castle AL, Elashoff MR;

PI WPI; 2002-241625/29.

XX Predicting toxic effects of compounds or the progression of these toxic

PT effects by determining the changes in gene expression in tissues or

PT cells exposed to the toxin and comparing these to gene expression in

PT unexposed tissues or cells -

XX Claim 1; Seq ID No 1424; 239pp; English.

XX The invention relates to methods for predicting toxic effects of

CC compounds or the progression of these toxic effects by determining the

CC global changes in gene expression in tissues or cells exposed to the

CC toxin and comparing these to gene expression in unexposed tissues or

CC cells. Also included are methods of predicting at least one toxic

CC effect of a compound or progression of a toxic effect, preferably the

CC hepatotoxicity of a compound, comprising detecting the level of

CC expression in a tissue or cell sample exposed to the compound of two or

CC more genes listed in the specification, where differential expression of

CC the genes is indicative of at least one toxic effect or progression.

CC The method can also be used to identify an agent which modulates the

CC toxic response and predict cellular pathways that a compound modulates

CC in a cell. The methods utilise a set of at least two probes (on a solid

CC support in kit form) where each of the probes comprises a sequence that

CC specifically hybridises to a gene listed in the specification, a computer

CC system comprising a database containing information identifying the

CC expression level in a tissue or cell sample exposed to a hepatotoxin of a

CC set of genes comprising at least two genes listed in the specification,

CC and a user interface to view the information used to present information,

CC identifying the expression level in a tissue or cell of at least one gene

CC listed in the specification. The method is useful for elucidating global

CC changes in gene expression and for identifying toxicity markers in

CC tissues or cell exposed to a known toxin. The genes may be used as

CC toxicity markers in drug screening and toxicity assays. The genes and

CC gene expression information may be used as diagnostic markers for the

CC prediction or identification of the physiological state of tissue or cell

CC sample that has been exposed to a compound or agent. Hepatotoxicity

CC is characterised by centrilobular necrosis and steatosis. The present

CC sequence is an expressed sequence tag (EST) or cDNA derived from a gene

CC which is differentially expressed in response to a hepatotoxic agent.

XX

SQ Sequence 4254 BP; 1188 A; 934 C; 1066 G; 1066 T; 0 other;

Alignment Scores:

Pred. No.: 9.17e-52 Length: 4254

Score: 467.00 Matches: 90

Percent Similarity: 91.06% Conservative: 22

Best Local Similarity: 73.17% Mismatches: 11

Query Match: 70.86% Indels: 0

DB: 24 Gaps: 0

US-09-873-409-3 (1-131) x ABK63517 (1-4254)

Qy 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21

Db 1456 ATTGACGGACGACATCAGGACCATCAATGTGAGGTATCTGCGGAAATCATTTGGGGTG 1515

Qy 22 ValSerGlnGluProValIleuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41

Db 1516 GTGAGTCAGGAACCGTGTCTTTTCCACCGATTCGCGAAACATTGCTATGCGCGA 1575

Qy 42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61

Db 1576 GAAACGTCACCATGATGAGATGAGAAAGCTGTCAAGGAAGCAATGCCATGACTTC 1635

Qy 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81

Db 1636 ATCATGAAACTGCCCCACAAATTTGACACCTGTTGTGTGAGAGAGGGCGCAGCTGACT 1695

Qy 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101

Db 1696 GGGGGACAGAAACAGAGGATGCCATTGCCGGGCCCTGTGTCGCAAGCAACCCCAAGATCTTT 1755

Qy 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121

Db 1756 TTGTTGGATGAGCCACGTCAGCCTTGACACACAGAAAGCAAGCCGTGTTCAGGCCGCT 1815

Qy 122 LeuGluLys 124

Db 1816 CTGGATAAG 1824

RESULT 12

AAZ52047

ID AAZ52047 standard; cDNA; 4369 BP.

AC AAZ52047;

XX 18-JUL-2000 (first entry)

DE Rat multidrug resistance protein 1a cDNA.

KW Multidrug resistance protein 1a; mdria; multi-specific drug transporter;

KW drug formulation; formulation excipient; compound design; inflammation;

KW pharmacokinetic; oral absorption; cancer; cardiovascular disease;

KW central nervous system disorder; auto-immune disease; kidney disease; ss.

OS Rattus rattus.

XX Key Location/Qualifiers

FT CDS 352..4170

FT /*tag= a

FT /product= "Rat multidrug resistance protein"

XX WO200015650-A1.

XX 23-MAR-2000.

XX 10-SEP-1999; 99WO-US20770.

XX 17-SEP-1998; 98US-0156800.

XX 09-DEC-1998; 98US-0208809.

XX (SMIK) SMITHKLINE BEECHAM CORP.

XX Brun KA, Chenery RJ, Ellens H, Feild JA, Yue L;

XX WPI; 2000-271372/23.

DR P-PSDB; AAY70596.

XX Isolated rat mdria polynucleotides and polypeptides, useful in assays

PT to provide information on drug formulation, selection of formulation

PT excipients and compound design -

XX Claim 2; Page 23-24; 33pp; English.

XX The present cDNA sequence encodes rat multidrug resistance

CC protein 1a (mdria). This protein is a member of the multi-specific drug

CC transporters family. Mdria is used in assays to provide information on

CC drug formulation, selection of formulation excipients and compound

CC design. They are used in cell based, membrane based, binding or other
CC assays to provide information that may enhance drug formulation. This
CC invention further relates to the generation of in vivo and in vitro
CC comparison data to predict oral absorption and pharmacokinetics. This
CC enables the selection of drugs with optimal pharmacokinetics, i.e. good
CC oral bioavailability, brain penetration, plasma half life, and minimum
CC drug interaction. Transgenic and knock-out animals created using rat
CC mdrla provides an insight into treating and preventing human diseases
CC including cancer, inflammation, cardiovascular disease, central nervous
CC system disorders, auto-immune and kidney disease.
XX
SQ Sequence 4369 BP; 1241 A; 973 C; 1096 G; 1059 T; 0 other;

Alignment Scores:
Pred. No.: 9,51e-52 Length: 4369
Score: 467.00 Matches: 90
Percent Similarity: 91.06% Conservativity: 22
Best Local Similarity: 73.17% Mismatches: 11
Query Match: 70.86% Indels: 0
DB: 21 Gaps: 0

US-09-873-409-3 (1-131) x AAZ52047 (1-4369)

QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
DB 1684 ATCGACGGACAGACATCAGACATCAATGTGAGGTATCTGCGGGAATCATTTGGGGTG 1743
QY 22 ValSerGlnGluProValLeuPheGlyThrIleSerAsnAsnIleLysTyrGlyArg 41
DB 1744 GTGAGTCAGAACCCCGTGTGTTGCCACCAATTCGCCGAAACATTCGTATGCGCGA 1803
QY 42 AspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
DB 1804 GAAACGTCACCATGATGATAGAGAAAGCTGTCAAGGAAGCCATGCTATGATTTTC 1863
QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
DB 1864 ATCATGAACTGCCCAAAATTTGACACCTCGTGTGTGAGAGAGGGCGCAGCTGAGT 1923
QY 82 GlyGlyGlnLysGlnArgIleAlaAlaArgAlaLeuValArgAsnProLysIleLeu 101
DB 1924 GGGGACAGAAACAGAGGATCGCATTCGCCGGCCCTGTCGCAACCCCAAGATCCTT 1983
QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaA 121
DB 1984 TTGTTGATGAGGCCACGTGAGCCTTGACACAGAAAGCGGCTGTTAGCGCGCT 2043

QY 122 LeuGluLys 124

DB 2044 CTGGATAG 2052

RESULT 13

AAZ52048

ID AAZ52048 standard; cDNA; 4425 BP.

XX AC AAZ52048;

XX 18-JUL-2000 (first entry)

XX Rat multidrug resistance protein 1a cDNA derived from EST sequences.

XX Multidrug resistance protein 1a; mdrla; multi-specific drug transporter;
XX drug formulation; formulation excipient; compound design; inflammation;
XX pharmacokinetic; oral absorption; cancer; cardiovascular disease;
XX central nervous system disorder; auto-immune disease; kidney disease;
XX EST; expressed sequence tag; ss.

XX Rattus rattus.

XX Key Location/Qualifiers

XX CDS 352..4170

XX /*tag= a

XX /product= "Rat multidrug resistance protein"

XX WO200015650-A1.
XX 23-MAR-2000.
XX 10-SEP-1999; 99WO-US20770.
XX 17-SEP-1998; 98US-0156800.
XX 09-DEC-1998; 98US-0208809.
XX (SMIK) SMITHKLINE BEECHAM CORP.
XX Brun KA, Chenery RJ, Ellens H, Feild JA, Yue L;
XX WPI: 2000-271372/23.
XX P-PSDB; AAY70597.
XX Isolated rat mdrla polynucleotides and polypeptides, useful in assays
XX to provide information on drug formulation, selection of formulation
XX excipients and compound design -
XX Claim 11; Page 27-29; 33pp; English.

XX The present cDNA sequence encodes rat multidrug resistance protein 1a
XX (mdrla). This cDNA is derived from EST (expressed sequence tag)
XX sequences. Mdr1a is used in assays to provide information on
XX drug formulation, selection of formulation excipients and compound
XX design. They are used in cell based, membrane based, binding or other
XX assays to provide information that may enhance drug formulation. This
XX invention further relates to the generation of in vivo and in vitro
XX comparison data to predict oral absorption and pharmacokinetics. This
XX enables the selection of drugs with optimal pharmacokinetics, i.e. good
XX oral bioavailability, brain penetration, plasma half life, and minimum
XX drug interaction. Transgenic and knock-out animals created using rat
XX mdrla provides an insight into treating and preventing human diseases
XX including cancer, inflammation, cardiovascular disease, central nervous
XX system disorders, auto-immune and kidney disease.

XX SQ Sequence 4425 BP; 1250 A; 992 C; 1113 G; 1067 T; 2 U; 1 other;

Alignment Scores:

Pred. No.: 9,68e-52 Length: 4425
Score: 467.00 Matches: 90
Percent Similarity: 91.06% Conservativity: 22
Best Local Similarity: 73.17% Mismatches: 11
Query Match: 70.86% Indels: 0
DB: 21 Gaps: 0

US-09-873-409-3 (1-131) x AAZ52048 (1-4425)

QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
DB 1684 ATCGACGGACAGACATCAGGACCATCAATGTGAGGTATCTGCGGGAATCATTTGGGGTG 1743
QY 22 ValSerGlnGluProValLeuPheGlyThrIleSerAsnAsnIleLysTyrGlyArg 41
DB 1744 GTGAGTCAGAACCCCGTGTGTTGCCACCAATTCGCCGAAACATTCGTATGCGCGA 1803
QY 42 AspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
DB 1804 GAAACGTCACCATGATGATAGAGAAAGCTGTCAAGGAAGCCATGCTATGATTTTC 1863
QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
DB 1864 ATCATGAACTGCCCAAAATTTGACACCTCGTGTGTGAGAGAGGGCGCAGCTGAGT 1923
QY 82 GlyGlyGlnLysGlnArgIleAlaAlaArgAlaLeuValArgAsnProLysIleLeu 101
DB 1924 GGGGACAGAAACAGAGGATCGCATTCGCCGGCCCTGTCGCAACCCCAAGATCCTT 1983
QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaA 121
DB 1984 TTGTTGATGAGGCCACGTGAGCCTTGACACAGAAAGCGGCTGTTAGCGCGCT 2043

QY 122 LeuGluLys 124
 DB 2044 CTGGATAAG 2052

RESULT 14
 ID AAF86127 standard; cDNA; 4186 BP.
 XX
 AC AAF86127;
 DT 25-JUN-2001 (first entry)
 XX
 DE Cynomologous monkey P-glycoprotein cDNA.
 XX
 KW Cynomologous monkey; P-glycoprotein; PGP; multidrug transporter; MDR1;
 KW efflux pump; ss.
 XX
 OS Macaca fascicularis.
 XX
 FH Key Location/Qualifiers
 FT CDS 100..3942
 FT /tag= a
 FT /product= "PGP"
 FT /note= "P-glycoprotein"
 XX
 PN WO200123565-A1.
 PD 05-APR-2001.
 XX
 PF 28-SEP-2000; 2000WO-US26592.
 PR 28-SEP-1999; 99US-0156921.
 PR 12-OCT-1999; 99US-0158818.
 XX
 PA (GENT-) GENTEST CORP.
 XX
 PI Stocker PJ, Steimel-Crespi DT, Crespi CL;
 XX
 DR WPI; 2001-316136/33.
 DR P-PSDB; AAB81064.
 XX

Novel isolated nucleic acid encoding cynomologous monkey P-glycoprotein (PGP) and homologous PGP polypeptides are useful for predicting cell bioavailability of compound and increasing PGP transporter activity in cell -

Example 1; Page 51-57; 84pp; English.

This invention relates to a polynucleotide sequence encoding a cynomologous monkey P-glycoprotein (PGP), and an allelic variant of the PGP protein. PGP, also known as multidrug transporter, MDR1 is a member of the ABC transporter superfamily. The enzyme serves as an efflux pump exporting small molecules across the cell membrane. The invention includes a cynomologous monkey (Macaca fascicularis) PGP coding sequence and protein, and also that of an allelic variant. The PGP polynucleotide sequence is useful for increasing PGP transporter activity in a cell. Antisense sequences of the cDNA are useful for inhibiting PGP transport activity in a mammalian cell. They may also be used for increasing the bioavailability of a drug. The present sequence represents cDNA encoding cynomologous monkey P-glycoprotein.

SQ Sequence 4186 BP; 1226 A; 799 C; 1039 G; 1122 T; 0 other;

Alignment Scores:
 Pred. No.: 1.66e-51 Length: 4186
 Score: 465.00 Matches: 91
 Percent Similarity: 88.62% Conservative: 18
 Best Local Similarity: 73.98% Mismatches: 14
 Query Match: 70.56% Indels: 0
 DB: 22 Gaps: 0

us-09-873-409-3 (1-131) x AAF86127 (1-4186)

QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
 DB 1456 GTTGATGGACAGGATATTAGGACCATAAACGTAAGTTTCTACGGGAATCATCGGTGTG 1515

QY 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
 DB 1516 GTGAGTGCAGGAACCTGTATTGTTGCCACCGATAGCTGAAACACATTCCGCTATGCTGT 1575

QY 42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
 DB 1576 GAAGATGTCACCATGATGAGATTGAGAAAGCTGTCAGGAAGCCATGCCCTATGACTTT 1635

QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
 DB 1636 ATCATGAAACTCCCTCAGAAATTTGACACCTGTTGGAGAGAGAGGGGCCCGAGCTGAGT 1695

QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
 DB 1696 GGTGGCGCAGAACGACGAGGATGCCATTGCACGTGCCCTGGTTCGCAACCCCAAGATCCTC 1755

QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
 DB 1756 CTGCTGGACGAGCCACGTCACGCTTGGACACAGAAAGTGAAGCAGTGGTTCAGGTGGCT 1815

QY 122 LeuGluLys 124
 DB 1816 CTGGATAAG 1824

RESULT 15
 ID AAF86128 standard; cDNA; 4195 BP.
 AC AAF86128;
 XX
 DT 25-JUN-2001 (first entry)
 XX
 DE Cynomologous monkey P-glycoprotein cDNA variant 1.
 XX
 KW Cynomologous monkey; P-glycoprotein; PGP; multidrug transporter; MDR1;
 KW efflux pump; ss.
 XX
 OS Macaca fascicularis.
 FH Key Location/Qualifiers
 FT CDS 100..3951
 FT /tag= a
 FT /product= "PGP"
 FT /note= "P-glycoprotein"
 FT /tag= b
 FT /note= "Insertion of 9 nucleotides relative to PGP allelic variant AAF86127"
 XX
 PN WO200123565-A1.
 PD 05-APR-2001.
 PF 28-SEP-2000; 2000WO-US26592.
 XX
 PR 28-SEP-1999; 99US-0156921.
 PR 12-OCT-1999; 99US-0158818.
 XX
 PA (GENT-) GENTEST CORP.
 XX
 PI Stocker PJ, Steimel-Crespi DT, Crespi CL;
 XX
 DR WPI; 2001-316136/33.
 DR P-PSDB; AAB81065.
 XX
 PT Novel isolated nucleic acid encoding cynomologous monkey P-glycoprotein (PGP) and homologous PGP polypeptides are useful for predicting PT bioavailability of compound and increasing PGP transporter activity in


```
PT cell -
XX
PS
XX Example 1: Page 59-65; 84pp; English.
XX
CC This invention relates to a polynucleotide sequence encoding a
CC cynomolgus monkey P-glycoprotein (PGP), and an allelic variant of the
CC PGP protein. PGP, also known as multidrug transporter, MDR1 is a member
CC of the ABC transporter superfamily. The enzyme serves as an efflux pump
CC exporting small molecules across the cell membrane. The invention
CC includes a cynomolgus monkey (Macaca fascicularis) PGP coding sequence
CC and protein, and also that of an allelic variant. The PGP polynucleotide
CC sequence is useful for increasing PGP transporter activity in a cell.
CC Antisense sequences of the cDNA are useful for inhibiting PGP transport
CC activity in a mammalian cell. They may also be used for increasing the
CC bioavailability of a drug. The present sequence represents cDNA encoding
CC cynomolgus monkey P-glycoprotein. This sequence contains a 9 nucleotide
CC insert compared to the PGP allelic variant given in AAF86127.
XX
SQ Sequence 4195 BP; 1230 A; 801 C; 1039 G; 1125 T; 0 other;

Alignment Scores:
Pred. No.: 1,66e-51 Length: 4195
Score: 465.00 Matches: 91
Percent Similarity: 88.62% Conservative: 18
Best Local Similarity: 73.98% Mismatches: 14
Query Match: 70.56% Indels: 0
DB: 22 Gaps: 0

US-09-873-409-3 (1-131) x AAF86128 (1-4195)
QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
DB 1465 GTTGATGGACAGCATATTAGSACCATAAAGCTAAGGTTTCTACGGGAAATCATCGGTGTG 1524
QY 22 ValSerGlnGluProValLeuPheGlyThrIleSerAsnAsnIleLysTyrGlyArg 41
DB 1525 GTGAGTCAGGAACCTGTATTGTTTGGCCACCACGATAGCTGAAAAACATTGCTATGGTCGT 1584
QY 42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
DB 1585 GAAGATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAATGCTATGACTTT 1644
QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
DB 1645 ATCATGAACCTGCTCAGAAATTGACACCTTGTTGGACAGAGGGGCCACAGCTGAGT 1704
QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
DB 1705 GGTGGGAGAGAGAGAGATCGCCATTGCACGTGCCCTGGTTGCGCAACCCCAAGATCCTC 1764
QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
DB 1765 CTGCTGACGAGGCCACGTGAGCCTTGGACACAGAAAGTGAAGCAGTGTTCAGGTGGCT 1824
QY 122 LeuGluLys 124
DB 1825 CTGGATAAG 1833
```

Search completed: March 30, 2003, 03:08:18
Job time : 142.031 secs

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 30, 2003, 01:24:25 ; Search time 1770.98 Seconds
(without alignments)
2152.747 Million cell updates/sec

Title: US-09-873-409-3
Perfect score: 659
Sequence: 1 MYDENDIRALNVRHYRDHIG.....SESKSAVQAALCKDTPRYSF 131

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DRV=xlh
-Q/cgn2_1/USPTO_spool/US09873409/runat_27032003.115418.19228/app_query.fasta_1.7544
-DB=GenEmbl -QMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdd -LIST=45
-DOCALIGN=200 -THR SCORE=pcp -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09873409 @CGN 1.1 54887 @runat_27032003.115418.19228 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DBLOP=6 -DELEXT=7

Database :

1:	gb.ba.*	GenEmbl.*
2:	gb.htg.*	
3:	gb.in.*	
4:	gb.cm.*	
5:	gb.ov.*	
6:	gb.pat.*	
7:	gb.ph.*	
8:	gb.pl.*	
9:	gb.pr.*	
10:	gb.ro.*	
11:	gb.sts.*	
12:	gb.sy.*	
13:	gb.un.*	
14:	gb.vi.*	
15:	em.ba.*	
16:	em.fun.*	
17:	em.hum.*	
18:	em.in.*	
19:	em.mu.*	
20:	em.om.*	
21:	em.or.*	
22:	em.ov.*	
23:	em.pat.*	
24:	em.ph.*	
25:	em.pl.*	
26:	em.ro.*	
27:	em.sts.*	
28:	em.un.*	

RESULT 1

ALIGNMENTS

29:	em.vi.*	
30:	em.htg.hum.*	
31:	em.htg.inv.*	
32:	em.htg.other.*	
33:	em.htg.mus.*	
34:	em.htg.pln.*	
35:	em.htg.rod.*	
36:	em.htg.nam.*	
37:	em.htg.vrt.*	
38:	em.sy.*	
39:	em.htgo.hum.*	
40:	em.htgo.mus.*	
41:	em.htgo.other.*	

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match	Length	DB	ID	Description
1	659	100.0	1175	6	AX339029	Sequence	AX339029 Sequence
2	659	100.0	1940	6	AX339034	Sequence	AX339034 Sequence
3	659	100.0	2021	6	AX339033	Sequence	AX339033 Sequence
4	619	93.9	2856	6	AX339028	Sequence	AX339028 Sequence
5	619	93.9	3177	6	AX339030	Sequence	AX339030 Sequence
6	619	93.9	3621	6	AX339032	Sequence	AX339032 Sequence
7	619	93.9	3699	6	AX478104	Sequence	AX478104 Sequence
8	619	93.9	3702	6	AX339031	Sequence	AX339031 Sequence
9	474	71.9	3489	4	AB029153	Felis cat	AB029153 Felis cat
10	472	71.6	3888	10	AY082609	Rattus no	AY082609 Rattus no
11	472	71.6	4233	6	AR123273	Sequence	AR123273 Sequence
12	469	71.2	2752	5	AF099732	Fundulus	AF099732 Fundulus
13	469	71.2	3334	4	AF419568	Canis fam	AF419568 Canis fam
14	469	71.2	4018	5	GGA9799	Gallus ga	AJ009799 Gallus ga
15	469	71.2	4045	12	AF269224	Synthetic	AF269224 Synthetic
16	469	71.2	4279	6	AX105057	Sequence	AX105057 Sequence
17	469	71.2	4279	6	AX105078	Sequence	AX105078 Sequence
18	469	71.2	4279	6	AX105080	Sequence	AX105080 Sequence
19	469	71.2	4279	6	AX105082	Sequence	AX105082 Sequence
20	468	71.0	4189	6	AX322791	Sequence	AX322791 Sequence
21	468	71.0	4298	10	MUSMDR	Mouse multi	M14757 Mouse multi
22	468	71.0	4356	10	MUSMDRAA	Mouse multi	M30697 Mouse multi
23	467	70.9	3858	4	OAU78609	Ovis aries	U78609 Ovis aries
24	467	70.9	3987	10	CRUPGP1185	Chinese ham	M5254 Chinese ham
25	467	70.9	4254	6	AX401748	Sequence	AX401748 Sequence
26	467	70.9	4254	10	RATMDRM	Rat mdr mrn	M81855 Rat mdr mrn
27	467	70.9	4280	10	CRUPGPII	C.griseus p	M60041 C.griseus p
28	467	70.9	4296	10	CRUPGPI	Chinese ham	M60040 C.griseus p
29	467	70.9	4304	10	CRUPGP1165	Chinese ham	M59253 Chinese ham
30	467	70.9	4323	10	AF286167	Rattus no	AF286167 Rattus no
31	467	70.9	4927	10	AF257746	Rattus no	AF257746 Rattus no
32	465	70.6	2005	4	AB066299	Canis fam	AB066299 Canis fam
33	465	70.6	4186	6	AX108654	Sequence	AX108654 Sequence
34	465	70.6	4195	6	AX108656	Sequence	AX108656 Sequence
35	465	70.6	4317	4	AF045016	Canis fam	AF045016 Canis fam
36	465	70.6	4317	6	AX105059	Sequence	AX105059 Sequence
37	465	70.6	4390	5	XLU17608	Xenopus lae	U17608 Xenopus lae
38	462	70.1	3682	10	MUSPGLY1A	Mouse phosp	M24417 Mouse phosp
39	462	70.1	4788	6	AX322793	Sequence	AX322793 Sequence
40	462	70.1	4924	10	MUSMDR1A	Mouse p-gly	M33581 Mouse p-gly
41	461	70.0	3840	6	AX481416	Sequence	AX481416 Sequence
42	461	70.0	3860	6	AX322787	Sequence	AX322787 Sequence
43	461	70.0	3860	6	AX322789	Sequence	AX322789 Sequence
44	461	70.0	3988	6	AX024454	Sequence	AX024454 Sequence
45	461	70.0	4192	9	AF016535	Homo sapi	AF016535 Homo sapi

AX339029
LOCUS AX339029 1175 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 11 from Patent WO0194400.
ACCESSION AX339029
VERSION AX339029.1 GI:18129121
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Frank, M.H. and Sayegh, M.H.
TITLE A gene encoding a multidrug resistance human p-glycoprotein
homologue on chromosome 7p15-21 and uses thereof
JOURNAL Patent: WO 0194400-A 11 13-DEC-2001;
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES
source Location/Qualifiers
1. .1175
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 337 a 238 c 258 g 342 t
ORIGIN
Alignment Scores:
Pred. No.: 4.24e-60 Length: 1175
Score: 659.00 Matches: 131
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-873-409-3 (1-131) x AX339029 (1-1175)
Qy 1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
Db 418 ATGGTGGATGAGAAATGACATCAGAGCTTTAAATGTGCGCATTTATCGAGACCATATTGGA 477
Qy 21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
Db 478 GTGGTTAGTCAAGAGCTTTCTTAATAATTTAAATGTCGGACCATCATGTAACAATATCAAGTATGGA 537
Qy 41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
Db 538 CGAGATGATGTGACTGATGAAGAGATGAGAGAGCAGCAAGGGAAGCAAAATCGCTATGAT 597
Qy 61 PheIleMetGluPheProAsnLysPheAsnThrIleuValGlyGluLysGlyAlaGlnMet 80
Db 598 TTTATCATGGAGTTTCTTAATAATTTAAATGTCGGACCATCATGTAACAATATCAAGTATGGA 657
Qy 81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100
Db 658 AGTGGAGGCGCAAGAACACAGAGATCGCAATTCCTGTCCTTAGTTCGAAACCCCAAGATT 717
Qy 101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120
Db 718 CTGATTTTATAGTAGGCTACGCTCGCCCTGGATTTCAGAAAGCAAGTCAGCTGTTCAAGCT 777
Qy 121 AlaLeuGluLysAspThrProArgTyrSerPhe 131
Db 778 GCACCTGGAGAGGATACCCCGAGGTATTCATTT 810
RESULT 2
AX339034
LOCUS AX339034 1940 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 16 from Patent WO0194400.
ACCESSION AX339034
VERSION AX339034.1 GI:18129126
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE

AUTHORS Frank, M.H. and Sayegh, M.H.
TITLE A gene encoding a multidrug resistance human p-glycoprotein
homologue on chromosome 7p15-21 and uses thereof
JOURNAL Patent: WO 0194400-A 16 13-DEC-2001;
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES
source Location/Qualifiers
1. .1940
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 530 a 394 c 448 g 568 t
ORIGIN
Alignment Scores:
Pred. No.: 7.66e-60 Length: 1940
Score: 659.00 Matches: 131
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-873-409-3 (1-131) x AX339034 (1-1940)
Qy 1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
Db 1183 ATGGTGGATGAGAAATGACATCAGAGCTTTAAATGTGCGCATTTATCGAGACCATATTGGA 1242
Qy 21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
Db 1243 GTGGTTAGTCAAGAGCTTTCTTAATAATTTAAATGTCGGACCATCATGTAACAATATCAAGTATGGA 1302
Qy 41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
Db 1303 CGAGATGATGTGACTGATGAAGAGATGAGAGAGCAGCAAGGGAAGCAAAATCGCTATGAT 1362
Qy 61 PheIleMetGluPheProAsnLysPheAsnThrIleuValGlyGluLysGlyAlaGlnMet 80
Db 1363 TTTATCATGGAGTTTCTTAATAATTTAAATGTCGGACCATCATGTAACAATATCAAGTATGGA 1422
Qy 81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100
Db 1423 AGTGGAGGCGCAAGAACACAGAGATCGCAATTCCTGTCCTTAGTTCGAAACCCCAAGATT 1482
Qy 101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120
Db 1483 CTGATTTTATAGTAGGCTACGCTCGCCCTGGATTTCAGAAAGCAAGTCAGCTGTTCAAGCT 1542
Qy 121 AlaLeuGluLysAspThrProArgTyrSerPhe 131
Db 1543 GCACCTGGAGAGGATACCCCGAGGTATTCATTT 1575
RESULT 3
AX339033
LOCUS AX339033 2021 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 15 from Patent WO0194400.
ACCESSION AX339033
VERSION AX339033.1 GI:18129125
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Frank, M.H. and Sayegh, M.H.
TITLE A gene encoding a multidrug resistance human p-glycoprotein
homologue on chromosome 7p15-21 and uses thereof
JOURNAL Patent: WO 0194400-A 15 13-DEC-2001;
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES
source Location/Qualifiers
1. .2021
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 547 a 406 c 462 g 605 t
ORIGIN

Alignment Scores:					
Pred. No.:	8.03e-60	Length:	2021		
Score:	659.00	Matches:	131		
Percent Similarity:	100.00%	Conservative:	0		
Best Local Similarity:	100.00%	Mismatches:	0		
Query Match:	100.00%	Indels:	0		
DB:	6	Gaps:	0		
US-09-873-409-3 (1-131) x AX339033 (1-2021)					
QY	1	MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyArgAspHisileGly	20		
Db	1264	ATGGTGGATGAGAAATGACATCAGAGCTTTAAATGTGCGCATTCGAGACCATTATGA	1323		
QY	21	ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAnileysTyGly	40		
Db	1324	GTGGTTAGTCAAGAGCGCTGTTTTGTTCCGGGCCACCATCAGTAACAATATCAAGTATGA	1383		
QY	41	ArgAspAspValThrAspGluMetGluArgAlaAalaArgGluAlaAsnAlaTyRasp	60		
Db	1384	CGAGATGATGTGACTGATGAAGAGATGGAGAGCAGCAAGGGAAGCAAAATGCATATGAT	1443		
QY	61	PheileMetGluPheProAnlysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet	80		
Db	1444	TTTATCATGGAGTTTCTTAATAAATTTAATACATTTGTTAGGGAAGAAGAGCTCAAATG	1503		
QY	81	SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysile	100		
Db	1504	AGTGAGGGCAGAAAACAGAGGATCGCAATTCCTGCTGCTTAGTTTGGAAACCCCAAGATT	1563		
QY	101	LeulleAspGluAlaThrSerAlaLeuAspSerGluSerIySerAlaValGlnAla	120		
Db	1564	CTGATTTTAGATGAGGCTAGCTCTGCCCTGGATTCAGAAAGCAAGTCAGCTGTTCAAGCT	1623		
QY	121	AlaLeuGluLysAspThrProArgTyRserPhe	131		
Db	1624	GCACTGGAGAAGGATACCCCGCAGGTATTCATTT	1656		
RESULT 4					
LOCUS	AX339028	2856 bp	DNA	linear	PAT 09-JAN-2002
DEFINITION	Sequence 10 from Patent WO0194400.				
ACCESSION	AX339028				
VERSION	AX339028.1	GI:18129120			
KEYWORDS	human.				
SOURCE	Homo sapiens				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 Frank, M.H. and Sayegh, M.H. A gene encoding a multidrug resistance human p-glycoprotein homologue on chromosome 7p15-21 and uses thereof Patent: WO 0194400-A 10 13-DEC-2001; THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)				
AUTHORS					
TITLE					
JOURNAL					
FEATURES	Location/Qualifiers				
source	1..2856				
/organism="Homo sapiens"	/db_xref="taxon:9606"				
BASE COUNT	888 a	540 c	652 g	776 t	
ORIGIN					
Alignment Scores:					
Pred. No.:	1.9e-55	Length:	2856		
Score:	619.00	Matches:	124		
Percent Similarity:	100.00%	Conservative:	0		
Best Local Similarity:	100.00%	Mismatches:	0		
Query Match:	93.93%	Indels:	0		
DB:	6	Gaps:	0		
US-09-873-409-3 (1-131) x AX339028 (1-2856)					
QY	1	MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyArgAspHisileGly	20		
Db	739	ATGGTGGATGAGAAATGACATCAGAGCTTTAAATGTGCGCATTCGAGACCATTATGA	798		
QY	21	ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAnileysTyGly	40		
Db	799	GTGGTTAGTCAAGAGCGCTGTTTTGTTCCGGGCCACCATCAGTAACAATATCAAGTATGA	858		
QY	41	ArgAspAspValThrAspGluMetGluArgAlaAalaArgGluAlaAsnAlaTyRasp	60		
Db	859	CGAGATGATGTGACTGATGAAGAGATGGAGAGCAGCAAGGGAAGCAAAATGCATATGAT	918		

Qy 61 PheileMetGluProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80
|
|
|
Db 919 TTTATCATGAGTTTCTTAATAATTAAATACATTTGGTAGGGAAAAAGAGCTCAAATG 978
|
|
|
Qy 81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100
|
|
|
Db 979 AGTGGAGGCGAGAAACAGAGATCGCAATTGCTCGTCCCTTAGTTTCGAAACCCCAAGATT 1038
|
|
|
Qy 101 LeuileLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120
|
|
|
Db 1039 CTGATTTTAGATGAGGTACGCTGCGCTGGATTTCAGAAACCAAGTCAGCTGTTCAAGCT 1098
|
|
|
Qy 121 AlaLeuGluLys 124
|
|
|
Db 1099 GCACCTGGAGAG 1110
|
|
|
RESULT 6
LOCUS AX339032 3621 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 14 from Patent WO0194400.
ACCESSION AX339032
VERSION AX339032.1 GI:18129124
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Frank,M.H. and Sayegh,M.H.
TITLE A gene encoding a multidrug resistance human p-glycoprotein
JOURNAL homologue on chromosome 7p15-21 and uses thereof
PATENT: WO 0194400-A 14 13-DEC-2001;
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES
source
1. .3621
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 1081 a 696 c 842 g 1002 t
ORIGIN
Alignment Scores:
Pred. No.: 2,51e-55 Length: 3621
Score: 619.00 Matches: 124
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 93.93% Indels: 0
DB: 6 Gaps: 0
US-09-873-409-3 (1-131) x AX339032 (1-3621)
Qy 1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
|
|
|
Db 1183 ATGTGGATGAGATGACATCAGAGCTTTAAATGTGGCCATTATCGAGACCAATATTGGA 1242
|
|
|
Qy 21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
|
|
|
Db 1243 GTGGTTAGTCAAGAGCTGTTTGTTCGGACCCACCATCAGTAACATATCAAGATATGGA 1302
|
|
|
Qy 41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
|
|
|
Db 1303 CGAGATGATGTGACTGATGAAGAGATGAGAGAGCAGCAAGGGAAGCAAAATCGCTATGAT 1362
|
|
|
Qy 61 PheileMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80
|
|
|
Db 1363 TTTATCATGAGATTCTCTTAATAATTAAATCATTTGGTAGGGAAAAAGAGCTCAAATG 1422
|
|
|
Qy 81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100
|
|
|
Db 1423 AGTGGAGGCGAGAAACAGAGATCGCAATTGCTCGTCCCTTAGTTTCGAAACCCCAAGATT 1482
|
|
|
Qy 101 LeuileLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120
|
|
|

Db 1483 CTGATTTTAGATGAGCTAGCTCTGCCCTGGATTTCAGAAAGCAAGTCAGCTGTTCAAGCT 1542
|
|
|
Qy 121 AlaLeuGluLys 124
|
|
|
Db 1543 GCACCTGGAGAG 1554
|
|
|
RESULT 7
LOCUS AX478104 3699 bp DNA linear PAT 12-AUG-2002
DEFINITION Sequence 34 from Patent WO0240541.
ACCESSION AX478104
VERSION AX478104.1 GI:22217064
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Tang,Y.T., Yue,H., Nguyen,D.B., Hafalia,A.J., Elliott,V.S., Lu,Y.,
Walia,N.K., Yao,M.G., Baughn,M.R., Gandhi,A.R., Ding,L.,
Sanjanwala,M., Ramkumar,J., Arvizu,C., Gietzen,K.J., Lal,P.G.,
Azinzai,Y., Khan,F.A., Thangavelu,K., Thornton,M., Lu,D.A.,
Tribouley,C.M., Warren,B.A., Ison,C.H., Das,D., Raumann,B.E.,
Policky,J.L. and Kearney,L.
TITLE Transporters and ion channels
JOURNAL Patent: WO 0240541-A 34 23-MAY-2002;
Incyte Genomics, Inc. (US)
FEATURES
source
1. .3699
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="Incyte ID No: 7472030CB1"
BASE COUNT 1116 a 707 c 860 g 1016 t
ORIGIN
Alignment Scores:
Pred. No.: 2,58e-55 Length: 3699
Score: 619.00 Matches: 124
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 93.93% Indels: 0
DB: 6 Gaps: 0
US-09-873-409-3 (1-131) x AX478104 (1-3699)
Qy 1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
|
|
|
Db 1339 ATGTGGATGAGATGACATCAGAGCTTTAAATGTGGCCATTATCGAGACCAATATTGGA 1398
|
|
|
Qy 21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
|
|
|
Db 1399 GTGGTTAGTCAAGAGCTGTTTGTTCGGACCCACCATCAGTAACATATCAAGATATGGA 1458
|
|
|
Qy 41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
|
|
|
Db 1459 CGAGATGATGTGACTGATGAAGAGATGAGAGAGCAGCAAGGGAAGCAAAATCGCTATGAT 1518
|
|
|
Qy 61 PheileMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80
|
|
|
Db 1519 TTTATCATGAGATTCTCTTAATAATTAAATCATTTGGTAGGGAAAAAGAGCTCAAATG 1578
|
|
|
Qy 81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100
|
|
|
Db 1579 AGTGGAGGCGAGAAACAGAGATCGCAATTGCTCGTCCCTTAGTTTCGAAACCCCAAGATT 1638
|
|
|
Qy 101 LeuileLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120
|
|
|
Db 1639 CTGATTTTAGATGAGGTACGCTCTGCCCTGGATTTCAGAAAGCAAGTCAGCTGTTCAAGCT 1698
|
|
|
Qy 121 AlaLeuGluLys 124
|
|
|
Db 1699 GCACCTGGAGAG 1710
|
|
|

REFERENCE										
AUTHORS	Okai, Y., Nakamura, N., Matsushiro, H., Kato, H., Setoguchi, A., Yazawa, M., Okuda, M., Watari, T., Hasegawa, A. and Tsujimoto, H.									
TITLE	Molecular analysis of multidrug resistance in feline lymphoma cells									
JOURNAL	Am. J. Vet. Res. 61 (9), 1122-1127 (2000)									
MEDLINE	20431258									
REFERENCE	2 (bases 1 to 3489)									
AUTHORS	Tsujimoto, H. and Okai, Y.									
TITLE	Direct Submission									
JOURNAL	Submitted (05-JUN-1999) Hajime Tsujimoto, The University of Tokyo, Department of Veterinary Internal Medicine, The University of Tokyo, Bunkyo-ku, Tokyo 113-8657, Japan									
FEATURES	(E-mail: atsueji@hongo.ecc.u-tokyo.ac.jp, Tel:+81-3-5841-8004, Fax:+81-3-5841-8178)									
source	Location/Qualifiers									
	1. .3489									
	/organism="Felis catus"									
	/db_xref="taxon:9685"									
	/cell_line="adriamycin-resistant subline of FT-1"									
	/cell_type="feline lymphoma cell line"									
CDS	<1. .>3489									
	/note="multi-drug resistance related mRNA"									
	/codon_start=1									
	/protein_id="BAA87071.1"									
	/db_xref="GI:6472653"									
	/translation="HGAALPLMLVFGDMTDSFANAGISRNLTTLNITGESIVNDSYF									
	INRLEEMTTVAYYSGIGAGVLAAYIQVSPFCLAAGRIILKIRKQFPFHAMRQEVG									
	WPDVHDVGLNTRLDYVSKINEGIDKIMFQSMATFFIGFVGTGKWLTLVL									
	ALSPVLGLSAAIWAKILASPTDKELLAYAKAGAAVEVLAAITVIAFGQKKELERY									
	NKNLEAKRIGIKKAITANISIGIAFLIYASYALAFWYGTSLVLSHEYSIGQVLTVP									
	FSVLIGAFSVGQASPIEAFANARGAAVEIFKIDNKPISIDSYKNGHKPDNINKGLE									
	FKNVHFSYPSRKEVKILGLNLKVGSGGTVALVNGSGCGKSTVQLMORLYDPTGMV									
	SIDGQDIRTINVYRLREIIGVVSQEPVLFAITIAENIRYGRENVIMEETKAVKEANA									
	YDFIWLKPNKFTLVGERGAOLSGGQKORIAARALVRNPKILLDEATLSALDESEA									
	VYQVLDKARKGRTTVVAHRLSTIRNADVIAGFDGVIKEGNHDELMKESGIVFKL									
	VTMQTRNEIELENAYVIESISEIDALEMSPKDSGSLNRRSTRKSIHASQODRKLG									
	TKENLDERVPVPSFWRLIKNTTEPVPYVFGICAIINGQLQPAFVSILRIGVETR									
	VEDPETKQNSNIFSLFLGIISSITPFLQGTFGKAGEILTCLRITVYMRSLRQ									
	DYSWDFDPKNTGALTTLANDAAQVGAIGSLAIITQNIANLGTGIIISLIYQGL									
	TLLLAIVPIIAIAGVEMKLSGOALKDKKELEGKAVATEIENFRVTVSLTREQK									
	FEYMYAQSLOVPYRNSLRKAHIFGISFSTQAMYPYSYAGCFRFGALVAHEPMDQD									
	VLLVFSALFAGVAVQVSSFPADYAKAVSAAHVIMITEKIPLDISTSYGELMPNL									
	EGNVTFNEMFNPTRPDIPVLQGLSLEVKGTLLALVGGSGCGKSTVOLLERFYDP									
	MAGTVLIDGKEIKHLNVOLWRHMGIVSQEPLFDICIGENIAYGNSRVWSQEEIVR									
	AAKEANIHPIETLPKYNTRVGDKGTQLSGGQKORIAARALVRQPRILLDEATSA									
BASE COUNT	1015 a	716 c	847 g	911 t						
ORIGIN	LDTESEKVVQOEALDKAR"									
Alignment Scores:										
Pred. No.:	3.95e-40	Length:	3489							
Score:	474.00	Matches:	92							
Percent Similarity:	90.24%	Conservative:	19							
Best Local Similarity:	74.80%	Mismatches:	12							
Query Match:	71.93%	Indels:	0							
DB:	4	Gaps:	0							
US-09-873-409-3 (1-131) x AB029153 (1-3489)										
Qy	2	ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal	21							
Db	1180	ATTGATGGACAGACATTTAGGACCAATAATGTGAGATATCTTCGGGAAATCAATTGGTGTG	1239							
Qy	22	ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAenIleLysTyrGlyArg	41							
Db	1240	GTGAGTCAGAACTGTGTGTTCCTACCAATAGCTGAAACATCTCGATGCGCGT	1299							
Qy	42	AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAenAlaTyrAspPhe	61							
Db	1300	GAAATGTCCACCTGAGGAGGATGAGAAAGCTGTCAAGGAAGCAATGCTATGACTTT	1359							
Qy	62	IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer	81							

RESULT 8										
AX339031										
LOCUS	AX339031 linear PAT 09-JAN-2002									
DEFINITION	Sequence 13 from Patent WO0194400.									
ACCESSION	AX339031									
VERSION	AX339031.1 GI:18129123									
KEYWORDS										
SOURCE	human.									
ORGANISM	Homo sapiens									
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.									
REFERENCE	1									
AUTHORS	Frank, M.H. and Sayegh, M.H.									
TITLE	A gene encoding a multidrug resistance human p-glycoprotein homologue on chromosome 7p15-21 and uses thereof									
JOURNAL	Patent: WO 0194400-A 13 13-DEC-2001; THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)									
FEATURES	Location/Qualifiers									
source	1. .3702									
	/organism="Homo sapiens"									
	/db_xref="taxon:9606"									
BASE COUNT	1098 a	708 c	856 g	1039 t	1 others					
ORIGIN										
Alignment Scores:										
Pred. No.:	2.58e-55	Length:	3702							
Score:	619.00	Matches:	124							
Percent Similarity:	100.00%	Conservative:	0							
Best Local Similarity:	100.00%	Mismatches:	0							
Query Match:	93.93%	Indels:	0							
DB:	6	Gaps:	0							
US-09-873-409-3 (1-131) x AX339031 (1-3702)										
Qy	1	MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly	20							
Db	1264	ATGGTGGATGAGATGACATCAGACGCTTTAAATGTGGCGCATATATCGAGACCATATGGA	1323							
Qy	21	ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAenIleLysTyrGly	40							
Db	1324	GTGGTTAGTCAGAGCTGTTTGTTCGGACCAACATCATGTAACATATCAAGTATGGA	1383							
Qy	41	ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAenAlaTyrAsp	60							
Db	1384	CGAGATGATGTGACTGATGAAGAGATGGAGAGCAGCAGCAAGCAATGCGTATGAT	1443							
Qy	61	PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet	80							
Db	1444	TTTATCATGGAGTTTCCATAATTAATTAATCATTTGGTAGGGGAAAAGAGCTCAAATG	1503							
Qy	81	SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle	100							
Db	1504	AGTGGAGGGCAGAAACAGAGATCGCAATTCGTCTGCTGCTTAGTTCGAAACCCCAAGATT	1563							
Qy	101	LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla	120							
Db	1564	CTGATTTAGATGAGGTACGCTACGCTCGCTCGATTCAGAAAGCAAGTCAGCTGTTCAAGCT	1623							
Qy	121	AlaLeuGluLys	124							
Db	1624	GCACCTGGAGAAG	1635							
RESULT 9										
AB029153										
LOCUS	AB029153 linear MAM 28-MAR-2002									
DEFINITION	Felis catus multi-drug resistance related mRNA, partial cds.									
ACCESSION	AB029153									
VERSION	AB029153.1 GI:6472652									
KEYWORDS										
SOURCE	Felis catus feline lymphoma cell line									
	cell_line:adriamycin-resistant subline of FT-1 cdna to mRNA.									
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.									
REFERENCE	1									
AUTHORS	Okai, Y., Nakamura, N., Matsushiro, H., Kato, H., Setoguchi, A., Yazawa, M., Okuda, M., Watari, T., Hasegawa, A. and Tsujimoto, H.									
TITLE	Molecular analysis of multidrug resistance in feline lymphoma cells									
JOURNAL	Am. J. Vet. Res. 61 (9), 1122-1127 (2000)									
MEDLINE	20431258									
REFERENCE	2 (bases 1 to 3489)									
AUTHORS	Tsujimoto, H. and Okai, Y.									
TITLE	Direct Submission									
JOURNAL	Submitted (05-JUN-1999) Hajime Tsujimoto, The University of Tokyo, Department of Veterinary Internal Medicine, The University of Tokyo, Bunkyo-ku, Tokyo 113-8657, Japan									
FEATURES	(E-mail: atsueji@hongo.ecc.u-tokyo.ac.jp, Tel:+81-3-5841-8004, Fax:+81-3-5841-8178)									
source	Location/Qualifiers									
	1. .3489									
	/organism="Felis catus"									
	/db_xref="taxon:9685"									
	/cell_line="adriamycin-resistant subline of FT-1"									
	/cell_type="feline lymphoma cell line"									
CDS	<1. .>3489									
	/note="multi-drug resistance related mRNA"									
	/codon_start=1									
	/protein_id="BAA87071.1"									
	/db_xref="GI:6472653"									
	/translation="HGAALPLMLVFGDMTDSFANAGISRNLTTLNITGESIVNDSYF									
	INRLEEMTTVAYYSGIGAGVLAAYIQVSPFCLAAGRIILKIRKQFPFHAMRQEVG									
	WPDVHDVGLNTRLDYVSKINEGIDKIMFQSMATFFIGFVGTGKWLTLVL									
	ALSPVLGLSAAIWAKILASPTDKELLAYAKAGAAVEVLAAITVIAFGQKKELERY									
	NKNLEAKRIGIKKAITANISIGIAFLIYASYALAFWYGTSLVLSHEYSIGQVLTVP									
	FSVLIGAFSVGQASPIEAFANARGAAVEIFKIDNKPISIDSYKNGHKPDNINKGLE									
	FKNVHFSYPSRKEVKILGLNLKVGSGGTVALVNGSGCGKSTVQLMORLYDPTGMV									
	SIDGQDIRTINVYRLREIIGVVSQEPVLFAITIAENIRYGRENVIMEETKAVKEANA									
	YDFIWLKPNKFTLVGERGAOLSGGQKORIAARALVRNPKILLDEATLSALDESEA									
	VYQVLDKARKGRTTVVAHRLSTIRNADVIAGFDGVIKEGNHDELMKESGIVFKL									
	VTMQTRNEIELENAYVIESISEIDALEMSPKDSGSLNRRSTRKSIHASQODRKLG									
	TKENLDERVPVPSFWRLIKNTTEPVPYVFGICAIINGQLQPAFVSILRIGVETR									
	VEDPETKQNSNIFSLFLGIISSITPFLQGTFGKAGEILTCLRITVYMRSLRQ									
	DYSWDFDPKNTGALTTLANDAAQVGAIGSLAIITQNIANLGTGIIISLIYQGL									
	TLLLAIVPIIAIAGVEMKLSGOALKDKKELEGKAVATEIENFRVTVSLTREQK									
	FEYMYAQSLOVPYRNSLRKAHIFGISFSTQAMYPYSYAGCFRFGALVAHEPMDQD									
	VLLVFSALFAGVAVQVSSFPADYAKAVSAAHVIMITEKIPLDISTSYGELMPNL									
	EGNVTFNEMFNPTRPDIPVLQGLSLEVKGTLLALVGGSGCGKSTVOLLERFYDP									
	MAGTVLIDGKEIKHLNVOLWRHMGIVSQEPLFDICIGENIAYGNSRVWSQEEIVR									
	AAKEANIHPIETLPKYNTRVGDKGTQLSGGQKORIAARALVRQPRILLDEATSA									
BASE COUNT	1015 a	716 c	847 g	911 t						
ORIGIN	LDTESEKVVQOEALDKAR"									
Alignment Scores:										
Pred. No.:	3.95e-40	Length:	3489							
Score:	474.00	Matches:	92							
Percent Similarity:	90.24%	Conservative:	19							
Best Local Similarity:	74.80%	Mismatches:	12							
Query Match:	71.93%	Indels:	0							
DB:	4	Gaps:	0							
US-09-873-409-3 (1-131) x AB029153 (1-3489)										
Qy	2	ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal	21							
Db	1180	ATTGATGGACAGACATTTAGGACCAATAATGTGAGATATCTTCGGGAAATCAATTGGTGTG	1239							
Qy	22	ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAenIleLysTyrGlyArg	41							
Db	1240	GTGAGTCAGAACTGTGTGTTCCTACCAATAGCTGAAACATCTCGATGCGCGT	1299							
Qy	42	AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAenAlaTyrAspPhe	61							
Db	1300	GAAATGTCCACCTGAGGAGGATGAGAAAGCTGTCAAGGAAGCAATGCTATGACTTT	1359							
Qy	62	IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer	81							

[illegible]


```

Db 1436 GTGAGTCAGAAACCCGTCGTTTGGCCACACGATTGCCGAAACATTCGCTATGCGCGA 1495
      |||
Qy 42  AspAspValThrAspGluMetGluArgAlaAalaArgGluAlaAalaMetSer 81
      |||
Db 1496 GAAACGTCACCATGATGAGATAGAGAAAGCTGTCAAGGAAGCAATGCTATGACTTC 1555
      |||
Qy 62  IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
      |||
Db 1556 ATCATGAATGCGCCACACAAATTAACACCTGTTGGTGAGAGAGGGCGCAGCTGAGT 1615
      |||
Qy 82  GlyGlyGlnLysGlnArgGluAlaAalaArgAlaAalaMetSerProLysIleLeu 101
      |||
Db 1616 GGGGAGACAGAAACAGAGGATCGCCATTCGCCGGGCGCTGTCTCCCAACCAAGATCCCT 1675
      |||
Qy 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAala 121
      |||
Db 1676 TTGTTCGATGAGCGCCACGTCAGCTTGGACACAGAAAGCGGCGCTGTTCAGGCGCT 1735
      |||
Qy 122 LeuGluLys 124
      |||
Db 1736 CTCGATAAG 1744
      |||

RESULT 12
AF099732
LOCUS
DEFINITION
  Fundulus heteroclitus multidrug resistance transporter homolog
  (mdr) mRNA, partial cds.
ACCESSION
  AF099732
VERSION
  AF099732.1 GI:4574223
KEYWORDS
  Fundulus heteroclitus.
  Fundulus heteroclitus
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
  Acanthomorpha; Acanthopterygii; Percormorpha; Atherinomorpha;
  Cyprinodontiformes; Fundulidae; Fundulus.
  1 (bases 1 to 2752)
  Cooper, P.S., Van Veld, P.A. and Reese, K.S.
  P-glycoprotein related sequences from the mummichog (Fundulus
  heteroclitus)
  Unpublished
  2 (bases 1 to 2752)
  Cooper, P.S., Van Veld, P.A. and Reese, K.S.
  Direct Submission
  Submitted (19-OCT-1998) School of Marine Science, Department of
  Environmental Sciences, Virginia Institute of Marine Science,
  College of William and Mary, Gloucester Point, VA 23062, USA
  Location/Qualifiers
    1..2752
      /organism="Fundulus heteroclitus"
      /db_xref="taxon:8078"
      /tissue_type="liver; intestine"
      <1..2752
        /gene="mdr"
        <1..2556
          /gene="mdr"
          /codon_start=1
          /product="multidrug resistance transporter homolog"
          /protein_id="AAD23956.1"
          /db_xref="GI:4574224"
          /translation="STTIQLLRFYDQSGVSIDGHDIRSLNRYLRGMIGVVSQEP
            ILFATIAENIRYRDVTEIEKAAKANAYDFIMLPDKPFTLVGRGTOMSGGQ
            KORIAAPALVRPKTILLDEATSDALDAESETVOAALDKVQRGRTLLVAHCLSTIR
            NADVIAGLKPKVGLTSHSELMKKGVYTLVTQTFQKADGEDENLAGEKSP
            HNNVIESPLLRKSTGVSFAASIGEKDKQEKEDKEDDEEDFPMVSIKVLRLN
            ASVEFFYVLGLCATINGAIQPLFAVLFSKIIITVFAEPDQTIIRQRANFSLMFVIG
            AAQVQASGVRLTAQNTANLGTGVLFAVYVWELTLLI LAVPVIALAVQVKML
            TGHAAEDKKELEKAGKATEAENIRTVASLTREPKFESLYENLVYKNSQKKAHV
            YGFTFSQAMLYFAACFRGMLIVGRMDVAVFLVISAVLFGAMVGEANSFA
            PNYAKAMASHMLLNKEPEIDNLSEGESPDMPDGNVSFEDVKFNTPSPDPVIL
            RGLNLRVKKGETLALVSGSSCGCKSTTIQLLRFYDPRDGRVVMDSIDVLRNLHLRS
  
```

```

OIGIVSQBPVLFDCALENIAYGDSNRSVTMEETEAARAKAANIHFNINELPQKYNTQA
GDKQTLGGGQRRIATAIARNPKVLLDEALSALDTESEKVVQDLDQSKGRTIC
IVWHRUSTIRNADRIAVFQGVVVEQTHOOLLAKGVYHMLVTLQHGTE"

BASE COUNT 661 a 718 c 791 g 582 t
ORIGIN

Alignment Scores:
Pred. No.: 1e-39 Length: 2752
Score: 469.00 Matches: 91
Percent Similarity: 89.43% Conservative: 19
Best Local Similarity: 73.98% Mismatches: 13
Query Match: 71.17% Indels: 0
DB: 5 Gaps: 0

US-09-873-409-3 (1-131) x AF099732 (1-2752)
Qy 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyArgAspHisIleGlyVal 21
      |||
Db 58 ATCGATGGGACGACATCCGCTCTTATGTGGCTACTGAGGGGAATGATCGGCGTG 117
      |||
Qy 22 ValSerGlnGluProValLeuPheGlyThrIleSerAsnIleLysTyGlyArg 41
      |||
Db 118 GTGAGCAGGAGCCCATCTCTTGGCCACCACCATCGCTGAGAACATCCGCTACGCGCA 177
      |||
Qy 42 AspAspValThrAspGluMetGluArgAlaAalaArgGluAlaAalaMetSer 61
      |||
Db 178 CCGACGTCGACGAGGAGGAGATCGAAAAAGCTGCCAAGGAAGCCCAATGCTTACGACTTC 237
      |||
Qy 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
      |||
Db 238 ATCATGAACCTTCTGTATAGTTTGAGACACTAGTCGGGACCGGAGAACCCAGATGAGC 297
      |||
Qy 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
      |||
Db 298 GGAGGACAGACAGAGGATTCGATCGCTCGAGCTTTGGTCCGAAAAACCCAAAAATCCTG 357
      |||
Qy 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAala 121
      |||
Db 358 CTGTTGACGAAAGCCACGCTGCTCCCTGGATGCTGAGAGTGAGACTATCGTTCAAGCTGCA 417
      |||
Qy 122 LeuGluLys 124
      |||
Db 418 CTCGATAAG 426
      |||

RESULT 13
CFA419568
LOCUS
DEFINITION
  Canis familiaris mRNA for abortive P-glycoprotein (p-gp gene).
ACCESSION
  AJ419568
VERSION
  AJ419568.1 GI:17385398
KEYWORDS
  P-glycoprotein; p-gp gene.
  dog.
  Canis familiaris
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
  1
  Roulet, A.
  Characterization of a MDR1a P-gp deficient dog in relation to
  ivermectin sensitivity
  Thesis (2002) Department of Biological Sciences, Universite de
  Toulouse (UPS III), Toulouse, France
  2 (bases 1 to 3934)
  Roulet, A.
  Direct Submission
  Submitted (07-NOV-2001) Roulet A., Pharmacology, Inra, 180 Chemin
  de Tournefeuille, 31931 Toulouse cedex, FRANCE
  Location/Qualifiers
    1..3934
      /organism="Canis familiaris"
      /db_xref="taxon:9615"
      /tissue_type="liver"
      /note="Ivermectin sensitive collies"
    7..282
  
```



```

CDS
/genes="p-gp"
7..282
/genes="p-gp"
/note="abortive product"
/codon_start=1
/evidence=experimental
/product="p-glycoprotein"
/protein_id="CAD12806.1"
/db_xref="GI:17385399"
/translation="MDPEGRGSAEFKNFKWKKSKNEKKKPKTSTVSTFAMPRYSN
WLDRLMLVGTWAAIHGAALPLMLVFGNMTALQMFQETKLFQL"
BASE COUNT 1144 a 815 c 964 g 1011 t
ORIGIN

Alignment Scores:
Pred. No.: 1 52e-39 Length: 3934
Score: 469.00 Matches: 90
Percent Similarity: 89.43% Conservative: 20
Best Local Similarity: 73.17% Mismatches: 13
Query Match: 71.17% Indels: 0
DB: 4 Gaps: 0

US-09-873-409-3 (1-131) x CFA419568 (1-3934)
Qy 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
Db 1365 ATTGATGCAGACGACATTAGGACCATTAATGTTAAGGCATCTTCGGGAAATTAAGTGGTGTG 1424
Qy 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
Db 1425 GTGAGTCAGAGCGCTGTGTGTTTGCACACGACGATAGCTGCTGCTGCTGCTGCTGCTGCTG 1484
Qy 42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
Db 1485 GAAATGTCACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1544
Qy 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
Db 1545 ATCATGAACTACCTAATAATTTGACACTCTGTTGGAGAGAGAGAGAGAGAGAGAGAGAG 1604
Qy 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
Db 1605 GGTGGACAGAAACAGAGATCGCCATTTGCTGGGCCCTGCTGCTGCTGCTGCTGCTGCTG 1664
Qy 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
Db 1665 CTGCTGATGAGGCAAGTCACTGCTGACACTGCTGACACTGCTGACACTGCTGACACTGCTG 1724
Qy 122 LeuGluLys 124
Db 1725 CTGGATAAG 1733

RESULT 14
GGA9799
LOCUS Gallus gallus mRNA for ABC transporter protein. VRT 03-MAR-1999
DEFINITION Gallus gallus mRNA for ABC transporter protein.
ACCESSION AJ009799
VERSION AJ009799.1 GI:3355756
KEYWORDS ABC transporter protein; cmdr1 gene; P-glycoprotein.
SOURCE Gallus gallus
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 4018)
Edelmann,H.M.L.
Direct Submission
Submitted (24-JUL-1998) Edelmann H.M.L., Institute of Molecular
Genetics, Bioenter and University of Vienna, Dr. Bohr-Gasse 9/2
A-1030 Vienna, A-1030 Vienna, AUSTRIA
2 (bases 1 to 4018)
Edelmann,H.M., Duchek,P., Rosenthal,F.E., Foger,N., Glackin,C.,
Kane,S.E. and Kuchler,K.

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
```

```

TITLE Cmdr1, a chicken P-glycoprotein, confers multidrug resistance and
interacts with estradiol
JOURNAL Biol. Chem. 380 (2), 231-241 (1999)
MEDLINE 99209805
PUBMED 10195430
FEATURES
source
Location/Qualifiers
1..4018
/organism="Gallus gallus"
/db_xref="taxon:9031"
/tissue_type="intestine"
109..3975
/genes="cmdr1"
109..3975
/genes="cmdr1"
/note="ABC transporter protein"
/codon_start=1
/product="p-glycoprotein"
/protein_id="CAA08835.1"
/db_xref="GI:3355757"
/db_xref="SPTREMBL:O93437"
/translation="MHSEDEKQKHTVDGNTYIANSNQDQPEDDEKKGKKKKKKPKQM
VSPALFRYSCDKLLMIFGSLAIAHGTSLPIAMIIFGDMTDSFVTSQMTNITGNS
SGLNSSADVENKLEEMTRYAYVYSAIAAALVAAVYIOTSFWTLAGQVKKIRKFF
HAIHQEIGCFDNDAGELNTRLDDVSKINEGIDGKIGLIQSEFTLTGTGFIQFIR
GKWLTLVAVSPVLGSLAALWAKLLTAFTDKQAYAKAGAAVEVLSAVRTVAPG
QOEIKRYHNLEDAKIGIRKAITSNISGAAPFLIYASVALAPFYGTTLILANEY
STGNLTVFVSLIGAFSIGOTAPSIEAFNARGAAYAFNIIIDNEPIDSYDAGHK
PDHIGLEFQNFVFPSPDVEILKGLNKVNCQGTVALVGGSCGKSTVQIQR
FYDPKEGTTIDGDLKSLNRYLREIIGVNVQVPLFATTIENIRYGREVDWTMEI
ERATKANAYDFIMKLPKFETVGERGAQSGGQKQRIATARALVHNPKLILLDEAT
SALDTESVSVQAALDKAREGRTVVVAHRLSTVNRADLIAVFESGVITEQGNHSQLI
EKGIYYKLVNMQTTIEDPSSEKSNVSVKRSQSNLDESLEKLRGSTRSMK
KPEPNDDEKSSPDELPVPSFLKMLKNEMPFYVAGTFCATVNGALQAPFSVI
PSEIIGIFSETDQKLVKSLNLSLLFLALGIIISPTFVQGFAGKAGAILTMKLRF
WAFKMLRQDMAPDDPKNSTCALTRLANDASOVKGATGVRALIAQNIANGTGII
LSLVYQWQITLLLAVPVIAVAGMIEMKLAGHAKDKIELEAKGKATAEINRT
VASLRKRKRFELMYGHELLVYRNSVKAHIFGFCSLSQAMFTYACCFRPGAYLV
VNGHTEYKTVFLVFSVAGVAGALQTSFAPDYAKAKISAHLFLNFRVPPIDSYR
EDGEKPEKFGNTRIKDYFNYPNRPVKILQGLNLAVEKGETLALVSSGGCGTIV
OLLERFYDPLSGEIVDDIDAKTNIOWLRSHIGIVSQEPILFDFTIENIAYGNSR
EVSHEIISAACAASIHFSIDSLPEKYNTRVGDKGTOLSGGOKQRIATARALIRKPOI
LLLDATSAIDTESSEKIVQEALDKAREGRTCVIAHRLSTIQNADKIAVIQNGKVIQ
GTHQQLLAEKGFYISLVNVQSGSCNN"
mat_peptide
112..3972
/genes="cmdr1"
/product="ABC transporter protein"
/note="p-glycoprotein"
BASE COUNT 1288 a 758 c 911 g 1061 t
ORIGIN

Alignment Scores:
Pred. No.: 1 56e-39 Length: 4018
Score: 469.00 Matches: 91
Percent Similarity: 89.43% Conservative: 19
Best Local Similarity: 73.98% Mismatches: 13
Query Match: 71.17% Indels: 0
DB: 5 Gaps: 0

US-09-873-409-3 (1-131) x GGA9799 (1-4018)
Qy 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
Db 1489 ATTGACGGGCGAGGATCTGAAGAGCCTGAATGTAAGATATCTGCGAGAGATTATTGGTGTG 1548
Qy 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
Db 1549 GTGAACGAGAGCGCGGTGTTGCTTACAACTATTGCAAAATATTCGTTATGCGCGT 1608
Qy 42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
Db 1609 GAGGATGTACCATGGAAGAAATTTGAAGAGCTACCAAGGAGCTAATGCTTATGATTTC 1668
Qy 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 30, 2003, 03:22:44 ; Search time 7277.11 Seconds
(without alignments)
2354.618 Million cell updates/sec

Title: US-09-873-409-4

Perfect score: 5330

Sequence: 1 MVISITKSLAYSAGAVA.....QELLNRDIYFKLVNAQSVQ 1058

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=xlh
-Q/cgn2_1/USPTO_spool/US09873409/runat_27032003_115420_19240/app_query.fasta_1.7544
-DB=EST -Qfmt=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pcp -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09873409 @CGN 1.1 30544 @runat_27032003_115420_19240 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_man:*
24: em_gss_mus:*
25: em_gss_oth:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1092	20.5	669	13	BM013981	BM013981 603639619
2	1051	19.7	1033	14	BM926413	BM926413 AGENCOURT
3	1008	18.9	943	9	AL520322	AL520322 AL520322
4	995	18.7	760	13	BM016204	BM016204 603642659
5	918	17.2	545	9	AL040762	AL040762 DRFPZP434C
6	903	16.9	929	14	BQ882401	BQ882401 AGENCOURT
7	856.5	16.1	948	13	BM471690	BM471690 AGENCOURT
8	834	15.6	998	14	BM904842	BM904842 AGENCOURT
9	780	14.6	2676	11	AK014319	AK014319 Mus muscu
10	771	14.5	894	12	BF584668	BF584668 602098406
11	770.5	14.5	1019	12	BG248052	BG248052 602359987
12	753	14.1	726	12	BG293345	BG293345 602390738
13	752.5	14.1	871	17	AZ682350	AZ682350 ENTXB16TF
14	751	14.1	944	12	BF796582	BF796582 602258463
15	743.5	13.9	886	17	BH139685	BH139685 ENTNA47TR
16	734.5	13.8	932	17	AZ670821	AZ670821 ENTUN69TF
17	728.5	13.7	947	17	AZ683753	AZ683753 ENTIL96TF
18	726.5	13.6	886	17	AZ540627	AZ540627 ENTQ18TF
19	723.5	13.6	880	17	AZ687805	AZ687805 ENTU52TF
20	721.5	13.5	939	14	BQ720763	BQ720763 AGENCOURT
21	720.5	13.5	897	17	AZ541090	AZ541090 ENTDS67TR
22	719	13.5	780	12	BG587938	BG587938 EST489713
23	711.5	13.3	913	17	BH155700	BH155700 ENTRO54TR
24	709.5	13.3	834	17	AZ548312	AZ548312 ENTFG07TR
25	708.5	13.3	853	17	AZ679807	AZ679807 ENTH167TR
26	705.5	13.2	823	17	AZ532602	AZ532602 ENTICR50TR
27	697.5	13.1	861	14	BQ717101	BQ717101 AGENCOURT
28	692.5	13.0	1341	11	AI108285	AI108285 Zea mays
29	690.5	13.0	999	17	BH154857	BH154857 ENTRO47TF
30	690	12.9	835	13	BJ349604	BJ349604 BJ349604
31	689.5	12.9	891	17	AZ682250	AZ682250 ENTMT68TR
32	688	12.9	1123	11	AY108485	AY108485 Zea mays
33	687	12.9	715	14	BQ869512	BQ869512 QGD6G09.Y
34	684	12.8	820	13	BJ356540	BJ356540
35	683	12.8	600	13	BM486593	BM486593 Pgsm2n.pko
36	682.5	12.8	919	17	AZ690701	AZ690701 ENTH168TR
37	680	12.8	926	14	BQ123477	BQ123477 EST609053
38	677.5	12.7	795	12	BF133560	BF133560 601900192
39	674.5	12.7	815	13	BI409057	BI409057 602961137
40	670	12.6	852	17	BH720383	BH720383 BOHZ294TF
41	669	12.6	547	13	BQ303711	BQ303711 BJ030711
42	667	12.5	937	17	BH150760	BH150760 ENTQ42TR
43	666	12.5	792	13	BJ352404	BJ352404
44	663	12.4	838	13	BJ355307	BJ355307 BJ355307
45	662	12.4	636	10	AV962688	AV962688 AV962688

ALIGNMENTS

RESULT 1
BM013981
LOCUS 603639619F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:5415556 5', linear EST 30-OCT-2001
DEFINITION mRNA sequence.
ACCESSION BM013981
VERSION BM013981.1 GI:16528335
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 669)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabs-r@mail.nih.gov

Tissue Procurement: DCTD/DTF

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM12060 row: f column: 05

High quality sequence stop: 664.

FEATURES

Location/Qualifiers

source

1..669
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5415556"
 /clone_lib="NIH_MGC_87"
 /tissue_types="mammary adenocarcinoma, cell line"
 /lab_hosts="DH10B (phage-resistant)"
 /note="Organ: breast; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 1.383 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH MGC Library."
 BASE COUNT 203 a 122 c 162 g 182 t

ORIGIN

Alignment Scores:

Pred. No.: 8,28e-115 Length: 669
 Score: 1092.00 Matches: 215
 Percent Similarity: 99.54% Conservative: 2
 Best Local Similarity: 98.62% Mismatches: 0
 Query Match: 20.49% Indels: 1
 DB: 13 Gaps: 0

US-09-873-409-4 (1-1058) x BM013981 (1-669)

Qy 128 AlaValPheSerValIleHisSerTyrCysIleGlyAlaAlaValProHisPhe 147
 Db :::: 76
 17 AGTGTTCCTTAATGTAATCATGACGAGTATTGTCATGGAGCAGCAGTCCCTCAGCTTT 76
 Qy 148 GluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLys 167
 Db 77 GAACTTCGCAATAGCCGAGGAGCTGCTTTTATATTTTCCAGGTATTGATAGAAA 136
 Qy 168 ProSerIleAspAsnPheSerThralaGlyTyrLysProGluSerIleGluGlyThrVal 187
 Db 137 CCCAGTATAGATAACTTTTCCACAGCTGGATATAAACCCTGAATCCATAGAGGAAGTGTG 196
 Qy 188 GluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGly 207
 Db 197 GAATTTAAATAATGTTCTTTTCAATATATCCATCAAGACCATCTATCAAGATTCTGAAAGGT 256
 Qy 208 LeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGly 227
 Db 257 CTGAATCTCAGAAATTAAGTCTGGAGAGACAGTCGCTTGGTCGTCTCAATGGCAGTGGG 316
 Qy 228 LysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMet 247
 Db 317 AAGAGTACGGTAGTCCAGCTTCTGCAGAGGTATATGATCCGGATGATGCTTTATCATG 376
 Qy 248 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 267
 Db 377 GTGGATGAGATGATCATCAGAGCTTTAAATGTGGCGCATTAATCCAGACCATATTGGAGTG 436
 Qy 268 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAnIleLysTyrGlyArg 287
 Db 437 GTTAGTCAAGACCCCTGTTTGTTCGGACCAACATCATGATCAATATCAAGTATGGACGA 496
 Qy 288 AspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 307
 Db 497 CATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAATGCGTATGATTTT 556

Qy 308 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyAlaGlnMetSer 327
 Db 557 ATCATGGAGTTTCTTAATAAATTTAATACATTCGTAGGGGAAAAGAGGCTCAATGAGT 616
 Qy 328 GlyGlyGlnLysGlnArg-IleAlaIleAlaArgAlaLeuValArgAsnPro 344
 Db 617 GGAGGGCAGAAACAGAGGCATCGCAATTGCTCGCTTAGTTTCGAAACCCC 668

RESULT 2

BM926413

LOCUS

DEFINITION BM926413 1033 bp mRNA linear EST 12-MAR-2002
 AGENCOURT 6600788 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5764845
 5', mRNA Sequence.

ACCESSION

BM926413

VERSION

BM926413.1

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 1033)

NIH-MGC http://mgi.nci.nih.gov/.

AUTHORS

National Institutes of Health, Mammalian Gene Collection (MGC)

TITLE

Unpublished (1999)

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgabs-r@mail.nih.gov

Tissue Procurement: Life Technologies, Inc.

cDNA Library Preparation: Life Technologies, Inc.

DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM12818 row: o column: 22

High quality sequence stop: 3

High quality sequence stop: 686.

Location/Qualifiers

1..1033

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5764845"

/clone_lib="NIH_MGC_114"

/lab_hosts="DH10B"

/note="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: EcoRV (destroyed); RNA source anonymous pool of 6
 male brains, age range 23-27 yo. Library is oligo-dT
 primed and directionally cloned (EcoRV site is destroyed
 upon cloning). Average insert size 1.5 kb, insert size
 range 1-3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 019. Note:
 this is a NIH MGC Library."

BASE COUNT 289 a 212 c 218 g 312 t

ORIGIN

Alignment Scores:

Pred. No.: 9,2e-110 Length: 1033
 Score: 1051.00 Matches: 221
 Percent Similarity: 94.49% Conservative: 2
 Best Local Similarity: 93.64% Mismatches: 13
 Query Match: 19.72% Indels: 3
 DB: 14 Gaps: 0

US-09-873-409-4 (1-1058) x BM926413 (1-1033)

Qy 129 ValPheSerValIleHisSerTyrCysIleGlyAlaAlaValProHisPheGlu 148
 Db 180 GTTTCCTTTAGTGTATCCATAGCAGTTATTCATGGACGAGCAGTCCCTCAGTTGAA 239
 Qy 149 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPro 168
 Db 240 ACCTTCGCAATAGCCCCGAGGAGCTGCTTTTATATTTTCCAGGTTCATTGATAGAAACCC 299


```

|||||
Db 785 ACCATCCAGATGACAGACTTAATAGTGTGTTTTCAGAAATGCGAGAGTCAAGGACATGG 844
|||||
Qy 1036 yThrHieGlnGluLeuLeuArgAsnArgAspIleTyxPhelYsLeuValAsnAlaGlnSe 1056
|||||
Db 845 CACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTTCAATGGTCAAGTGTCCAGGC 904
|||||
Qy 1056 r 1056
Db 905 T 905

RESULT 4
BM016204 603642659P1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:5418615 5',
LOCUS mRNA sequence.
DEFINITION BM016204
ACCESSION BM016204.1 GI:16530558
VERSION EST.
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 760)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-x@mail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L1AM12068 row: e column: 16
High quality sequence stop: 740.
FEATURES
Location/Qualifiers
source 1..760
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5418615"
/clone_lib="NIH_MGC_87"
/tissue_type="mammary adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: breast; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dr primed.
Average insert size 1.383 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT 229 a 150 c 187 g 193 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 1..55e-103 Length: 760
Score: 995.00 Matches: 203
Percent Similarity: 98.07% Conservatives: 0
Best Local Similarity: 98.07% Mismatches: 2
Query Match: 18.67% Indels: 2
DB: 13 Gaps: 0

US-09-873-409-4 (1-1058) x BM016204 (1-760)

Qy 166 LysLysProSerIleAspAsnPheserThrAlaGlyTyxLysProGluSerIleGluGly 185
Db 3 AAGAAACCCAGTATAGTAACCTTTTCCACAGCTGGATATAAACCCTGAATCCATGAAGGA 62
Qy 186 ThrValGluPheLysAsnValSerPheAsnTyxProSerArgProSerIleLysIleLeu 205
Db 63 ACTGTGGAATTTAAAGATGTTCTTCAATATATCATCAAGACCATCTATCAAGATCTG 122
Qy 206 LysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGly 225

```

```

|||||
Db 123 AAAGGTCTGAATCTCAAGATTAAGTCTGGAGAGACAGTCCGCTTGGTCTCAATGGC 182
|||||
Qy 226 SerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyxAspProAspAspGlyPhe 245
|||||
Db 183 AGTGGGAAGAGTACGGTAGTCCAGCTTCTCGAGAGGTATATATCCGGATGATGGCTTT 242
|||||
Qy 246 IleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyxArgAspHisIle 265
|||||
Db 243 ATCATGGTGGATGAGATGACATCAGAGCTTTAAATGTGGGCATTATCAGACCATATN 302
|||||
Qy 266 GlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyx 285
|||||
Db 303 GGAGTGGTTAGTCAAGAGCCCTGTTTGTTCGGACCAACCATCATAGTAACAATATCAGTAT 362
|||||
Qy 286 GlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyx 305
|||||
Db 363 GGACGAGATGATGTGACTGATGAAGAGATGGAGAGACAGCAAGGGGAAGCAATGCGTAT 422
|||||
Qy 306 AspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGln 325
|||||
Db 423 GATTTATCATGAGTTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 482
|||||
Qy 326 MetSerGlyGlnLysGlnArgIleAlaIle-AlaArgAlaLeuValArgAsnProLy 345
|||||
Db 483 ATGAGTGGAGGGCAGAAACACAGAGATCGCAATTTGGCTCGTCTAGTTCCGAAACCCCAA 542
|||||
Qy 345 sIleLeuIleLeu-AspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValG 365
|||||
Db 543 GATTTGATTTTTCAGATGAGGCTACGCTGCCCTGGATTGAGAAAGCAAGTCAAGCTGTC 602
|||||
Qy 365 InAlaAlaLeuGluLys 370
|||||
Db 603 AAGCTGCACCTGGAGAG 619

RESULT 5
AL040762 545 bp mRNA linear EST 29-FEB-2000
LOCUS DKF2p434C1815 r1 434 (synonym: htes3) Homo sapiens cDNA clone
DEFINITION DKF2p434C1815 5', mRNA sequence.
ACCESSION AL040762.1 GI:5409708
VERSION EST.
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 545)
AUTHORS Bloecker,H., Boecker,M., Brandt,P., Mewes,H.W., Gassenhuber,J. and
Wiemann,S.
TITLE EST (Bloecker, et al.)
JOURNAL Unpublished (1999)
COMMENT Contact: Bloecker H
MIPS
Am Klopferpitz 18a D-82152 Martinsried, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by GGF (National Research Center for Biotechnology Ltd.,
Braunschweig/Germany) within the cDNA sequencing consortium of the
German Genome Project.
s1 sequence also available.
This clone (DKF2p434C1815) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
Location/Qualifiers
source 1..545
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKF2p434C1815"
/clone_lib="434 (synonym: htes3)"
/tissue_type="testis"
/dev_stage="adult"

```



```

/lab_host="DH10B"
/notes=vector: psport1; Site_1: NotI; Site_2: SalI"
BASE COUNT 171 a 89 c 138 g 147 t
ORIGIN

Alignment Scores:
Pred. No.: 6,58e-95 Length: 545
Score: 918.00 Matches: 181
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 17.22% Indels: 0
DB: 9 Gaps: 0

US-09-873-409-4 (1-1058) x AL040762 (1-545)
QY 169 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 188
Db 3 AGTATAGATAACTTTCCACAGCTGGATATAACCTGAATCCATAGAGAACTGTGGAA 62
QY 189 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu 208
Db 63 TTTAAAATGTTCTTTCAATTATCCATCAAGACCATCTCAAGATTCTGAAAGGTCTG 122
QY 209 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 228
Db 123 AATCTCAGAAATTAAGTCTGGAGACAGTGCCTTGGTCTGCTCAATGGCAGTGGGAAG 182
QY 229 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 248
Db 183 AGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCGATGATGCTTTATCATGTG 242
QY 249 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal 268
Db 243 GATCAGAATGACATCAGAGCTTTAAATGTGCGCATTTATCGAGACCATATTTGGAGTGT 302
QY 269 SerGluProValLeuPheGlyThrIleSerAsnAsnIleLysTyrGlyArgAsp 288
Db 303 AGTCAAGAGCCTGTTTGTTCGGGACCACCATCAGTAAACAAATATCAAGTATGGACGAG 362
QY 289 AspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnLatyrAspPheIle 308
Db 363 GATGTGACTGATGAAGATGGAGAGCAGCAGGAGGAGCAATGCGTATGATTTTATC 422
QY 309 MetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGly 328
Db 423 ATGGAGTTTCTTAATTAATTAATACATTGTTAGGTAGGGGAAAAAGAGCTCAATGAGTGA 482
QY 329 GlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIle 348
Db 483 GGGCAGAAACAGAGGATCGCAATTGCTCGTCTAGTTCGAAACCCCAAGATTCTGATT 542
QY 349 Leu 349
Db 543 TTA 545

RESULT 6
LOCUS BQ882401
DEFINITION AGENCOURT_8627902 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:6291782
5', mRNA sequence.
ACCESSION BQ882401
VERSION BQ882401.1 GI:22274409
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 929)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
CONTACT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov

```

```

Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2492 row: c column: 15
High quality sequence stop: 677.
Location/Qualifiers
1. .929
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NIH_MGC_43"
/tissue_type="normal pigmented retinal epithelium"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: eye; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dr priming. Directionally-
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC library. |"
BASE COUNT 281 a 176 c 237 g 234 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 8,72e-93 Length: 929
Score: 903.00 Matches: 188
Percent Similarity: 73.02% Conservative: 61
Best Local Similarity: 55.13% Mismatches: 58
Query Match: 16.94% Indels: 34
DB: 14 Gaps: 3

US-09-873-409-4 (1-1058) x BQ882401 (1-929)
QY 12 AlaTyrSerLysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThr-ValI1 31
Db 1 GCGTATGCAAAAGCTGGAGCAGTAGCTGAAGAGTCTTGGCAGCAATTAGAACTGGTAT 60
QY 31 eAlaPheArgAlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTy 51
Db 61 TGCATTGGAGGACAAAGAAAGAACTTGAA----- 91
QY 51 rAlaTrpPheTyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrTh 71
Db 92 -----AGGTACAA 99
QY 71 rGlnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSe 91
Db 100 CAAAAATTTAGAGAAGCTTAAAGAAATTTGGATTAAGAAGCTATTACAGCAATATTTC 159
QY 91 rLeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyTh 111
Db 160 TATAGGTGCTGCTTCTCTGCTGATCTATGCTATCTATGCTCTGCTCTCTGCTATGGAC 219
QY 111 rSerLeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePh 131
Db 220 CACCTTGGTCTCTCAGGGGAA-----TATCTATTTGGACAAGTACTACTGTATTTCT 273
QY 131 eSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAl 151
Db 274 TTTCTGTTAATTTGGGGCTTTTAGTGTGGACAGGAGCTCTCCAGGCAATTGAAGCAATTGC 333
QY 151 aIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAs 171
Db 334 AAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATAATTAATAAAGCAAGTATTGA 393
QY 171 pAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAs 191
Db 394 CAGCTATTGCAAGAGTGGGCACAAACCAAGATATATTAAGGGAAATTTGGAATTCAGAA 453

```


QY 191 nValSerPheAsnTyrProSerArgProSerIleLeuLysGlyLeuAsnLeuAr 211
 Db 454 TGTTCACCTTACCTACCACTCGAAGAAGTTAAGATCTTGAAGGCTCTGAACCTGAA 513
 QY 211 gileLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrVa 231
 Db 514 GGTSCAGAGTGGGACAGCGTGGCCCTGGTTGGAAACAGTGGCTGTGGGAAGAGCACAA 573
 QY 231 lValGlnLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAs 251
 Db 574 AGTCACCTGATGCAGAGGCTCTATGACCCACAGAGGGGATGTCAGTGTGATGGACA 633
 QY 251 nAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGl 271
 Db 634 GGATATTAGACCAATAATGTAAGTTCTTACGGGAATCATTTGGTGTGTGAGTCAGGA 693
 QY 271 uProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspValTh 291
 Db 694 ACCTGTATTGTTCCACACAGATAGTGAATAACATTTCCCTATGGCCGTGAATAATGTCC 753
 QY 291 rAspGluGluMetGluArgAlaAlaArgGluAlaAsnAla-TyrAspPheIleMetGluP 311
 Db 754 CATGGATGAGATTGAGAAGCTGTCAGGAAGCAATGCCCTATGACTTTATCATGAAC 813
 QY 311 heProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGly---G 330
 Db 814 TGCCTCATTAATTTGACACCTGTTGGANNAAGAGGGCCCAATTCAGTGGTGGGCGAC 873
 QY 330 lnLysGlnArgIleAlaIleAlaArg-AlaLeuVal-ArgAsnProLysIleLeu 347
 Db 874 AACCAAAAGGATCTCCCTGGCAGCTGGCCCTGTTCCCAACCCCAAGATCCTC 928
 RESULT 7
 BM471690 948 bp mRNA linear EST 05-FEB-2002
 LOCUS AGENCOURT 6465349 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5539117
 DEFINITION 5', mRNA sequence.
 ACCESSION BM471690
 VERSION BM471690.1 GI:18520732
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 948)
 AUTHORS NIH-MGC http://mgs.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: ATCC/DCTP/DTF
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L14M12233 row: b column: 14
 High quality sequence stop: 569.
 Location/Qualifiers
 1. 948
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5539117"
 /clone_lib="NIH_MGC_72"
 /tissue_type="melanotic melanoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: skin; Vector: pCMV-SPORT6; Site: 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo; dt.
 Average insert size 2 Kb. Library constructed by Life
 Technologies."
 FEATURES
 source 285 a 187 c 194 g 282 t
 BASE COUNT

ORIGIN

Alignment Scores: 2,06e-87 Length: 948
 Pred. No.: 856.50 Matches: 186
 Score: 89.67% Conservative: 5
 Percent Similarity: 87.32% Mismatches: 14
 Best Local Similarity: 16.07% Indels: 8
 Query Match: 13
 DB: 1
 US-09-873-409-4 (1-1058) x BM471690 (1-948)
 QY 129 ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGlu 148
 Db 205 GTTTCTTTAGTGAATTCATAGCAGTATTCATTTGGAGCAGCAGTCCCTCACTTTGAA 264
 QY 149 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysPro 168
 Db 265 ACCTTCGCAATAGCCGAGGAGCTGCCTTTTCATATTTCCAGGTATTGTATAAGAACCC 324
 QY 169 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 188
 Db 325 AGTATAGATAACTTTTCCACAGCTGGATATAACCTGAATCCATAGAGGAAGCTGTGAA 384
 QY 189 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu 208
 Db 385 TTTAAAAATGTTCTTTCAATTATCCATCAAGACCATCTATCAAGATTCTGAAAGGTCTG 444
 QY 209 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 228
 Db 445 AATCTCAGAATTAAAGTCTGGAGAGCAGTCGCTTGGTGGTCTCAATGGCAGTGGGAAG 504
 QY 229 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 248
 Db 505 AGTACCGTAGTCAGCTTTCGACAGCTTATATGATCCGGATGATGCTTATCATGTTG 564
 QY 249 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIle-GlyValVa 268
 Db 565 GATGAGAATGACATCAGAGCTTTTAAATGTGCGCATTTATCGAGACCATATTTGGAGTGGT 624
 QY 268 lSerGlnGluPro--ValLeuPheGlyThrThrIleSerAsnAsnIleLysTyr-GlyAr 287
 Db 625 TAGTCAAGAGCCCTGTTGTTGTTGGGACCCACCATCAGTAACCAATATCAAGTATGGGACG 684
 QY 287 GAspAspValThrAspGluGlu-MetGluArgAlaAlaArgGluAla-AsnAlaTyrAsp 306
 Db 685 AGATATGGGACTGATGAACAAAATGGAANAAGCAACAGGGAACCAAGGCGAATGAA 744
 QY 307 PheIleMetGlu-PheProAsnLysPheAsnThrLeuVal---GlyGluLysGlyAlaGl 325
 Db 745 TTTATCATGAATTTCCCAAAAAAATTTAATACATTTGGAAGGGGAAAAAAGGACCTTCA 804
 QY 325 nMetSerGlyGlyGlnLysGlnArg 333
 Db 805 AGGAATGGAAGGCCCAAAACCCCAAG 829
 RESULT 8
 BM904842 998 bp mRNA linear EST 12-MAR-2002
 LOCUS AGENCOURT_6699581 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5557655
 DEFINITION 5', mRNA sequence.
 ACCESSION BM904842
 VERSION BM904842.1 GI:19355221
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 998)
 AUTHORS NIH-MGC http://mgs.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.

SOURCE	ORGANISM
Mus musculus (strain:C57BL/6J) 14, 17 days embryo head cDNA to mRNA, clone_lib:RIKEN full-length enriched mouse cDNA library clone:3222401P09.	Mus musculus
Eukaryotic; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
1	Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning	
Meth. Enzymol. 303, 19-44 (1999)	
95279253	
10349636	
2	Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes	
Genome Res. 10 (10), 1617-1630 (2000)	
20499374	
11042159	
3	Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kiteunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwaki,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer	
Genome Res. 10 (11), 1757-1771 (2000)	
20530913	
11076861	
4	Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y., Arakawa,T., Harada,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S., Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I., Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R.,
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
MEDLINE	
PUBMED	
REFERENCE	
AUTHORS	


```

Db 792 GTGCAAGAAATTCAGTATCATCTTCAAGCTTGTGTATCTCCAGGAGCTGCGCAGCTGCACC 851
Qy 623 AsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIle 642
Db 852 CAGGTGATGGTAGCTGGTGTCCCTGTCTATGCTGTCCCTCGCTTACCTCGTAGTGTCTG 911
Qy 643 LeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGly 662
Db 912 GCTGTGTACACACCCGCTCATGGAGTGGGCACCTGTATGGCTCAGGCTCCGAAAG 971
Qy 663 PheAlaAsnIleAspIleGlnGluLeuLeuHisAlaGlyIleIleAlaThrGluAlaLeu 682
Db 972 CTGTCTCGCAGCAGTGTGAGGAGCATGTCAGGCGCAACAGGTGTAGCAGATGAGGCTT 1031
Qy 683 GluAsnIleArgThrIleValSerLeuThrArgGluIleAlaPheGluGlnMetTyrGlu 702
Db 1032 GGCATGTTCGGACTGTGCGGGCTTCGCCATGAGAGAGGAGGAGGAGGAGGAGGAGGAG 1091
Qy 703 GluMetLeuGlnThrGlnHisArgAsnThrSerIleIleGlySerCys 722
Db 1092 GCAGAACTGGAG-----TCATGCTGTGTAAGCAGAGCAAGCACTGGCGCAGGGC 1139
Qy 723 TyrAlaPheSerHisAlaPheIleTyrPheAlaTyr-----AlaAlaGlyPheArg 739
Db 1140 ATCGCTTGTTCGAAGGCTCTCCCAACATCGCTTCAACTGTATGTCTTTGGGCAACCTG 1199
Qy 740 PheGlyAlaTyrIleGlnAlaGlyArgMetThrProGluGly-----MetPhe 756
Db 1200 TTCATTTGGGGCTCTCTGTGGCTGGACAGCAGCTGAAGGGGAGACCTCATGTCTCTTC 1259
Qy 757 IleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyIleThrLeuValLeuAla 776
Db 1260 CTGTGTGCTTCCAGACAGTACAGAGCTATGGCC-----AGCCTCTGTGCTCTGTTT 1313
Qy 777 ProGluTyrSerIleAlaIleSerGlyAlaAlaHisLeuPheAlaLeuLeuGlyLys 796
Db 1314 GGTCAAGTGTGTAGTGGCTGAGTGGCGGAGCCGAGCTCTTCGAATACATGCGCCCTGAGC 1373
Qy 797 ProAsnIleAspSerArgSerGlnGluGlyIleValSerProAspThrCysGluGlyAsnLeu 816
Db 1374 CCTGTATCCCATTTGACCGGGGGCTACTGCATCCCAACAGGACATTCCTGTGTTCATC 1433
Qy 817 GluPheArgGluValSerPheTyrProCysArgProAspValPheIleLeuArgGly 836
Db 1434 ACCTTCAAATGTCACCTTCAGTACCTCCCTCGACAGCTGGCTTCAATGTGCTCAAGGAC 1493
Qy 837 LeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGly 856
Db 1494 TTCACCTGAAGCTGCCCTCTGGCAAGATTGTGGCTCTGTGGGCCAGTCTGGGGGAGGA 1553
Qy 857 LysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeu 876
Db 1554 AAGACCAAGTGTGCTCTGTGCGAAGCTTCTATACCTGAAGCTGCTGCTCGGTGACG 1613
Qy 877 PheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGln---IleAla 895
Db 1614 TTGGATGGGATGACCTGCGAATCTCAACCTCTGCTGGCTCCGGGGCCAGGTCAATAGT 1673
Qy 896 IleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGly 915
Db 1674 TTCATCAGCAGGAGCCAGTCTCTGTTTGCACCAACCATCATGAGAGATATCCGATTGGG 1733
Qy 916 -----AspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAla 933
Db 1734 AAGCTGATGCTTCCGAT-----GAAGAGGTGTACACAGCTGCACGAGAGGCC 1781
Qy 934 AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLys 953
Db 1782 AATGCCACAGATTCATCAGCAGCTTCCCGATGGCTACAGCACTGTGGTTGGTGGACGG 1841
Qy 954 GlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGln 973

```

```

Db 1842 GGCACACACCTTCTGTGGTGGCCAGAGCAGCGCTAGCCATCGCAGCTGCCCTCATCAAG 1901
Qy 974 LysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 993
Db 1902 CAGCCACACAGTCTGATCTCTGGACGAGGCCACAGTGGCTAGATCGAGATCCGAGAGG 1961
Qy 994 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 1013
Db 1962 GTGTGTACAGAGGCCCTGACCGGGCCAGTGTGTGGCCGACCGCTGTGTGTATTTGCCAC 2021
Qy 1014 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys 1033
Db 2022 CGCTGTAGTACTTCTCGTGCAGCCCACTCCATTCATTTGTCGCAATGCCAAGTCTGT 2081
Qy 1034 GluGlnGlyThrHisGlnLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 1053
Db 2082 GAGCTGGGACCCACAGCAACTCTCTTAAAAAGGGGGCTCTATTTCAGAGCTTATCCGG 2141
Qy 1054 AlaGlnSerVal 1057
Db 2142 AGACAAACCCCTG 2153

RESULT 10
BF584668 894 bp mRNA linear EST 12-DEC-2000
LOCUS 602098406F1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4218385 5',
DEFINITION mRNA sequence.
ACCESSION BF584668
VERSION BF584668.1 GI:11658386
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 894)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9798 row: 1 column: 02
High quality sequence stop: 651.
FEATURES
source
1..894
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4218385"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: colon; Vector: pCMV-SPORT6; Site 1: Not I;
Site 2: SalI; Cloned unidirectionally. Primer: oligo dt.
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 263 a 203 c 242 g 185 t 1 others
ORIGIN
Alignment Scores: 1.33e-77 Length: 894
Pred. No.: 771.00 Matches: 171
Score: 79.72% Conservative: 57
Percent Similarity: 59.79% Mismatches: 48
Best Local Similarity: 14.47% Indels: 15
Query Match: 12 Gaps: 3
DB:
US-09-873-409-4 (1-1058) x BF584668 (1-894)

```



```

QY 129 ValPheSerValIleHisSerSerTyrcysileGlyAlaAlaValProHisPheGlu 148
Db 5 GTCTCTTTTCGGTGAAT-GGAGCATTCAGTGTGGACAGGATCTCCAAATATTGAA 63

QY 149 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysPro 168
Db 64 GCCTTCGCAATGACGAGGAGCAGCTTATGAAGCTTCAAAATATGATAAATAGCCC 123

QY 169 SerIleAspAsnPheSerThrAlaGlyTyrlsProGlnSerIleGluGlyThrValGlu 188
Db 124 AGTATACAGAGCTTCTCAAGAGATGGGCACAAACACACATACATCAAGGAAATCTCGAA 183

QY 189 PheLysAsnValSerPheAsnTyProSerArgProSerIleLysIleLeuLysGlyLeu 208
Db 184 TTTAAGATATATCACTTCAGTACCTACCTCGAAAGAGTTTCAGATCTTGAAGGGCTC 243

QY 209 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 228
Db 244 AATCTGAAGGTGAAGAGCGGACACAGCGTGGCCCTGTTGGCAACAGTGGCTGTGGAATA 303

QY 229 SerThrValGlnLeuLeuGlnArgLeuTyrlsProAspAspGlyPheIleMetVal 248
Db 304 AGCACAACTGTCCAGCTGATGCAAGGCTCTACGACCCCTAGATGGCATGGTCAGTATC 363

QY 249 AspGluAsnAspIleAlaAlaLeuAsnValArgHisTyrlsArgAspHisIleGlyValVal 268
Db 364 GACGACAGGACATCAGAACCATCAATGTAGGTATCTGAGGAGATCATGTGTGGTG 423

QY 269 SerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrlsGlyArgAsp 288
Db 424 AGTCAGAAACCTGTGCTGTGTGGCCACACGATCGCCGAGAACATTCGCTATGGCCGAA 483

QY 289 AspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrlsAspPheIle 308
Db 484 GATCTCACCATGGATGATGAGATTGAGAAAGCTGTCAAGAGCCCAATGCTATGACTTCATC 543

QY 309 MetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGlyAlaGlnMetSerGly 328
Db 544 ATGAACCTGCCCAACCATTTGACACCTGTGTGGTAGAGAGGGGCGAGCTGAGTGG 603

QY 329 GlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIle 348
Db 604 GGACAGAAACAGAGAAATCGC-ATTGCCGGGGCCCTGTGTCGCAATCCCAAGTGCCTTC- 661

QY 349 LeuAspGluAlaThrSerAlaLeuAspSerCluSerLysSerAlaValGlnAlaAla-Le 368
Db 662 GTGACCAAGCCACCTCAGC-CTGGATACAGAAAGTGAAGCTGGGGTTCAGGCGGCACCT 720

QY 368 uGluLysAlaSerLysGlyArgThrThrIle-----ValValAlaHisArgLe 384
Db 721 AGATAGGTACAAAGGCGGAGCACCATTTGGGAGCTCATCGCTGTTCGCTCGAAGCT 780

QY 384 uSerThrIleArgSerAlaAspLeuIleVal-----ThrLeuLysAspGlyMetLeuAl 402
Db 781 GACGCCATCGGGCGGAAGGGGGG-CTTGTGGGAACGACTCTGAAAGC-----TC 830

QY 402 aGluLysGlyAlaHis 407
Db 831 TGAAGAGGTACTCAT 846

RESULT 11
BG248052 1019 bp mRNA linear EST 13-FEB-2001
LOCUS BG248052 602359987F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:4488404 5',
DEFINITION mRNA sequence.
ACCESSION BG248052
VERSION BG248052.1 GI:12757867
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

```

REFERENCE 1 (bases 1 to 1019)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
plate: LLAM10334 row: n column: 21
High quality sequence stop: 650.
FEATURES
source
1..1019
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4488404"
/clone_lib="NCI_CGAP_Mam1"
/tissue_type="tumor, biopsy sample"
/dev_stage="10 months, virgin"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"
BASE COUNT 276 a 262 c 283 g 196 t
ORIGIN
Alignment Scores:
Pred. No.: 1.9e-77 Length: 1019
Score: 770.50 Matches: 154
Percent Similarity: 79.17% Conservative: 36
Best Local Similarity: 64.17% Mismatches: 47
Query Match: 14.46% Indels: 3
DB: 12 Gaps: 1
US-09-873-409-4 (1-1058) x BG248052 (1-1019)
QY 813 GluGlyAsnLeuGluPheArgGluValSerPheTyrcysArgProAspValPhe 832
Db 4 GAAGGAAATGTGCAATTTAGTGGAGTCGTGTCAACTATCCACCCGACCCAGCATCCCA 63
QY 833 IleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySer 852
Db 64 GTGCTTCAGGGGCTGAGCCTTGAGTGAAGAGGGCCAGACGCTGGCCCTGGTGGGCAGC 123
QY 853 SerGlyCysGlyLysSerThrSerValGlnLeuGlnArgLeuTyrlsAspProValGln 872
Db 124 AGTGGCTGCGGGAGAGACAGTGGTCCAGCTGCTCGAGCGCTTCTACGACCCCATGGCT 183
QY 873 GlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSer 892
Db 184 GGATCAGTGTCTTCTAGATGCAAGAAATAAAGCAACTGAATGTCCAGTGGCTCCGAGCA 243
QY 893 GlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIle 912
Db 244 CAGCTGGCATTTGTGTCCTCCAAAGAGCCCATCTCTTTGACTGTCAGCATCGCAGAGAACATT 303
QY 913 AlaTyrlsGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAla 932
Db 304 GCCTACGGAGACACAGCCGGGTCGTGCTTATGAGAGAGATTGTGAGGGGCGAGCCAGGAG 363
QY 933 AlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrlsAsnThrGlnValGlyLeu 952
Db 364 GCCAACATCCACCATTCATCGACTCGCTACCTGATAATAAACAACACAGTAGGAGAC 423
QY 953 LysGlyAlaGlnLeuSerGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeu 972
Db 424 AAAGGCACCTCAGCTGTGGGTGGGCAAGAGCAGCGCATGCCCATCGCACGCGCCCTGTC 483

```



```

Qy 973 GlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnSerGlu 992
      :::::|||||
Db 484 AGACAGCCTCACATTCTTCTGGACGAAGCAACATCAGCTCTGGATACAGAAAGTAA 543
      :::::|||||

Qy 993 LysValValGlnHisAlaLeuAspLysAlaAArgThrGlyArgThrCysLeuValValThr 1012
      :::::|||||
Db 544 AAGTTGTCCAGGAAGCGTGGCAAAAGCCAGGAAGGCGCACCTGCATTGTGATCGCT 603
      :::::|||||

Qy 1013 HisArgLeuSerAlaIleGlnAenAlaAspLeuIle-ValValLeuHis-AsnGlyLysI 1032
      :::::|||||
Db 604 CACCGTGTCCACCATCCAGAACGGGACTTGATCGTGGTGAATTCAGAACCGGAAG 663
      :::::|||||

Qy 1032 lelys---GluGlnGlyThrHisGlnGlnLeuArgAsnArgAspIleTyrPhe 1049
      :::::|||||
Db 664 TCAAGGGCAGCGGACCCCAACAGCAAGTGTGGCGCAGAGGGGCATCTACTTC 719
      :::::|||||

RESULT 12
BG293345 726 bp mRNA linear EST 21-FEB-2001
LOCUS 602390738F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:4502552 5',
DEFINITION mRNA sequence.
ACCESSION BG293345
VERSION BG293345.1 GI:13052943
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 726)
JOURNAL NIH-MGC http://mgc.nci.nih.gov/.
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: The Cepko Laboratory
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10371 row: 1 column: 09
High quality sequence stop: 658.
FEATURES
Location/Qualifiers
1..726
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:4502552"
/clone_lib="NIH MGC 94"
/tissue_type="retina"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 3.3 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC library."
BASE COUNT 189 a 220 c 190 g 127 t
ORIGIN

Alignment Scores:
Pred. No.: 1..11e-75 Length: 726
Score: 753.00 Matches: 152
Percent Similarity: 84.29% Conservative: 25
Best Local Similarity: 72.38% Mismatches: 30
Query Match: 14.13% Indels: 3
DB: 12 Gaps: 0

US-09-873-409-4 (1-1058) x BG293345 (1-726)

Qy 847 ValAlaPheValGlySerSerGlyCysGlyLysThrSerValGlnLeuLeuGlnArg 866
      :::::|||||
Db 2 CTGGCCCTGTGGGACAGTGGCTGGCGGAAGACACAGTGGTCCAGCTGCAGCGC 61
      :::::|||||

```

```

Qy 867 LeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGlnLeuAen 886
      :::::|||||
Db 62 TTCTATGACCCCATGGCTGGATCAGTGTCTTTAGATGGTCAAGAAAGCAAACTCAAT 121
      :::::|||||

Qy 887 ValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCys 906
      :::::|||||
Db 122 GTCCAGTGGCTCCGAGCTCAACTGGGCATTGTGCCAGAACCCCATCTCTTTGACTGC 181
      :::::|||||

Qy 907 SerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValProLeuAspGluIle 926
      :::::|||||
Db 182 AGCATCGCAGAGAACATCGCTATGGAGACAACAGCCGGTCTGTCCTCATGATGAGATT 241
      :::::|||||

Qy 927 LysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyr 946
      :::::|||||
Db 242 GTGAGGGCAGCCAGGAGGCCAACATCCACCCCTTCATCGAGACGCTGCCCAAAATAT 301
      :::::|||||

Qy 947 AsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlnLysGlnArgLeuAla 966
      :::::|||||
Db 302 AACACAAGAGTAGGAGACAGGGGAGCGAGCTCTCTGGGGGCCAGAGCAGAGATTGCC 361
      :::::|||||

Qy 967 IleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuAspGluAlaThrSerAla 986
      :::::|||||
Db 362 ATGCCCGGAGCCCTCATCAGACAGCTCGGGTCTTACTGCTGGATGAACCCAGCTCAGCT 421
      :::::|||||

Qy 987 LeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArg 1006
      :::::|||||
Db 422 CTGGATCTACTGAGAGTCAAAAGGTTGTCCAGGAAGCACTGGACAAGCCAGGAGCCCGC 481
      :::::|||||

Qy 1007 ThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAla-AspLeuIleValVa 1026
      :::::|||||
Db 482 ACCTGCATTTCGATCGCTCACCCTGTCCACATCCAGAACCGCGCACTTGATCGTGTGT 541
      :::::|||||

Qy 1026 lleuHisAsnGlyLysIleLysGluGlnGlyThr-HisGlnGluLeuLeuArgAsnArgA 1046
      :::::|||||
Db 542 GATTGAGACGGCAAGGTCAAGGACGACGCGACCCACCAGCAGCTGCTGGCGCAGAGG 601
      :::::|||||

Qy 1046 spileTyrPhe-LysLeuValAsn 1053
      :::::|||||
Db 602 GCATCTATTTCCTCAACTGTGTCAAC 625
      :::::|||||

RESULT 13
AZ682350 871 bp DNA linear GSS 14-DEC-2000
LOCUS ENTKB16TF Entamoeba histolytica Sheared DNA Entamoeba histolytica
DEFINITION genomic, DNA sequence.
ACCESSION AZ682350
VERSION AZ682350.1 GI:11819496
KEYWORDS GSS.
SOURCE Entamoeba histolytica.
ORGANISM Entamoeba histolytica.
REFERENCE 1 (bases 1 to 871)
AUTHORS Loftus,B., Van Aken,S. and Fraser,C.
TITLE Determination of clone end sequences from Entamoeba histolytica
HMI:IMSS sheared DNA library
JOURNAL Unpublished (2000)
COMMENT Contact: Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 3543
Email: bjloftus@igr.org
Clones are derived from the Entamoeba histolytica HMI:IMSS sheared
DNA library
Seq primer: M13-Forward
Class: shotgun
High quality sequence start: 16
High quality sequence stop: 860.
Location/Qualifiers
1..871
/organism="Entamoeba histolytica"
FEATURES
source

```


/strain="HMI:IMSS"
/db xref="taxon:5759"
/clone lib="Entamoeba histolytica Sheared DNA"
/notes="Vector: pHS1; Site 1: Bst I; Constructed at The
Institute for Genomic Research (TIGR), Rockville, MD.
Genomic DNA isolated from broth cultures of E. histolytica
using a method described by Clark and Diamond (Clark,
C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
method for isolate identification. Exp. Parasitol.
77:450). The DNA was mechanically sheared to give a
tight size distribution (~2 kb). The v + i method used for
the library construction is described in detail in Smith,
H.O. and Venter, J.C. (Making small insert libraries for
whole genome shotgun sequencing projects. In Genome
Sequencing: A Practical Approach, eds. M. Vaudin and B.
Barelli, Oxford University Press, 1999)."

BASE COUNT 349 a 117 c 174 g 231 t
ORIGIN

Alignment Scores:
Pred. No.: 1-72e-75 Length: 871
Score: 752.50 Matches: 151
Percent Similarity: 71.64% Conservative: 46
Best Local Similarity: 54.91% Mismatches: 75
Query Match: 14.12% Indels: 3
DB: 17 Gaps: 1

US-09-873-409-4 (1-1058) x AZ682350 (1-871)

Qy 153 AlaArgGlyAlaAlaPheHisIlePheGlnValIleAspIleAspSerIleAspAsn 172
Db 45 GCTAAGTTCGCGATTAAATGTTTATCAACAATTTGATAGATCCCGATATGATTGT 104
Qy 173 PheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnVal 192
Db 105 CAGTCTATTGGAGGTGAATGTCACACTGAGTGAATGAAATATATTAGATTGGAGATGT 164
Qy 193 SerPheAsnTyrProSerArgProSerIleLysIleLysGlyLeuAsnLeuArgIle 212
Db 165 CAATTTGTTTATCCACACAGACTGTCTCATCATGATTATTAAGACTTGACCTTGAATTT 224
Qy 213 LysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValVal 232
Db 225 AAGAAAGGACAAACAATTTGATTTAGTTGAGCATCAGGATGTGGAGTCAACTACTATT 284
Qy 233 GlnLeuLeuGlnArgLeuTyrAspProAspGlyPheIleMetValAspGluAsnAsp 252
Db 285 CAATTAATCCAAAGAAATTTATGATCCAAATGTTGGAAGAGTAACATTAGACGGAAAGAT 344
Qy 253 IleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluPro 272
Db 345 ATACGAGATTGATATATCAATGTTAGAAATCAATAGATTAGTTGGACAGACCA 404
Qy 273 ValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAsp----- 289
Db 405 GTGTTGTTTTCAGGAACAATTCGAGAAAATATTATGCTTGGAGCTAAAGAGGACCA 464
Qy 290 ValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMet 309
Db 465 CCAAGTGAAGAAGAGATGATTGAATGTCTAAAATGGCAATTCACATGACTTCTTCT 524
Qy 310 GluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyValAlaGlnMetSerGlyGly 329
Db 525 AAATCTCCAGAGGATATGACACAATAATTTGAGAAAAGGAGCAATTATTCAGGAGGA 584
Qy 330 GlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeu 349
Db 595 CAAAACAAAGAAATTCGAGTGCATTTGATGATTCGAAACCCATCTATTCTTCTT 644
Qy 350 AspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaLeuGlu 369
Db 645 GATGAAGCTACATCAGCCTTGATACACAAAGTGAAGAGATTGTACAAAGAGACCTTGA 704

Qy 370 LysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIleArgSer 389
Db 705 AAAGCATCTAAAGGAAGAACAAATATTATTAGCATACATAGACTAAACACTGTAGAAAT 764
Qy 390 AlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGlu 409
Db 765 GCAGATAAAATATGTCTATTTCATCAAGGAGAAATATTTCACACAGGAAACATCAAGAA 824
Qy 410 LeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln 424
Db 825 TTAATGATTGAAAGGACATATTATGATTAGTCAAAAGACAA 869
RESULT 14
BF796582
LOCUS 602258463F1 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:4341710 5',
DEFINITION mRNA sequence.
ACCESSION BF796582 944 bp mRNA linear EST 12-JAN-2001
VERSION BF796582.1 GI:12101636
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 944)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Louis Staudt, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM9955 row: n column: 15
High quality sequence start: 9
High quality sequence stop: 669.
Location/Qualifiers
1. .944
/organism="Homo sapiens"
/db xref="taxon:9606"
/clone="IMAGE:4341710"
/clone lib="NIH MGC 85"
/tissue_type="lymphoma, cell line"
/note="Organ: lymph; Vector: pCMV-Sport6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.867 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

BASE COUNT 265 a 241 c 237 g 201 t
ORIGIN

Alignment Scores:

Pred. No.: 2-94e-75 Length: 944
Score: 751.00 Matches: 166
Percent Similarity: 74.62% Conservative: 31
Best Local Similarity: 62.88% Mismatches: 60
Query Match: 14.09% Indels: 9
DB: 12 Gaps: 2

US-09-873-409-4 (1-1058) x BF796582 (1-944)

Qy 779 TyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuGluLysLysProAsn 798
Db 39 TATGCTAAAGCTAAGCTGTCTGCAGCCCACTTATTCTGCTTTGAAGACACCTCTG 98
Qy 799 IleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPhe 818
Db 99 ATTGACAGCTACAGTGAAGAGGGGCTGAGCCCTGATAAAATTTGAAGGAAATATAACATT 158

QY 819 ArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSer 838
 DB 159 AATGAAGTCGTGTCAACTATCCACCGAGCAACGTCAGTCGTTCAGGGCTGAGC 218
 QY 839 LeuSerIleGluArgGlyLysThrValAlaPheValIcysSerSerGlyCysGlyLysSer 858
 DB 219 CTGAGAGTGAAGAACGCCAGACACTAGCCCTGGTGGCGACAGTGGCTGTGGGAAGAGC 278
 QY 859 ThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAsp 878
 DB 279 ACGGTGTCAGCTCTCTGGAGCGGTTCTACGACCCCTTGGCGGGACAGTCTTCGAT 338
 QY 879 GlyValAspAlaLysGluLeuAenValGlnTrpLeuArgSerGlnIleAlaIleValPro 898
 DB 339 GGTCAAGAACCAAGAACTCAATGTCCAGTGGCTCAGAGCTCAACTCGGAATCGTGTCT 398
 QY 899 GlnGluProValLeuPheAenCysSerIleAlaGluAenIleAlaTyrGlyAspAenSer 918
 DB 399 CAGGAGCCTATCTCTATTGATCGAGCATTTGCCGAGAATATTGCCCTATGGAGACAACAGC 458
 QY 919 ArgValValProLeuAspGluIleLysGluAlaAlaAenAlaAenIleHisSerPhe 938
 DB 459 CGGTGTGTATCAGAGTGAATCGTCAGTGCAGCCAAAGCTGCCACATACATCTCTTC 518
 QY 939 IleGluGlyLeuProGluTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSer 958
 DB 519 ATCGAGAGCTTATCCCAAAATATGAAACAAAGAGTGGGAGATAAGGGAGCTCAGCTCTCA 578
 QY 959 GlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuLys-ProLysIleLe 978
 DB 579 GGAGGTCAACAAACAGAGGATGTGTTATGCCGAGCCCTCATCAGACAACCTCAAAATCCT 638
 QY 978 uLeuLeuAspGluAlaThrSerAla-LeuAspAenAspSerGluLysValVal- - - - -G 996
 DB 639 CCGTGTGTGATGAAGCTACATCAGCCTCTGGATACCTGACAGTGAACACCGTTGTCCAA 698
 QY 996 IHisAlaLeuAspLysAlaArgThrGlyArgThrCysLysLeuValValThrHisArgLeuS 1016
 DB 699 GAAGCCCTGTCCACAGCAGACAGAGCGCCGACCTGCTGTGATGCTCACCGCTGTC 758
 QY 1016 exAlaIleGlnAenAlaAspLeuIleValValLeuHisAenGlyLysIleLysGluGlnG 1036
 DB 759 AAC-ATCCAGAATGCAGACTT-ATAGTGTGTTCCTCCCAACGGGAGAGTCAAG- - - - - 808
 QY 1036 lyThrHis 1038
 DB 809 --ACACAT 814

RESULT 15
 BHI39685
 LOCUS ENTN477R Entamoeba histolytica Sheared DNA linear GSS 07-AUG-2001
 DEFINITION Genomic, DNA sequence.
 ACCESSION BHI39685
 VERSION BHI39685.1 GI:15098746
 KEYWORDS GSS.
 SOURCE Entamoeba histolytica.
 ORGANISM Entamoeba histolytica
 Eukaryota; Entamoebidae; Entamoeba.
 REFERENCE 1 (bases 1 to 886)
 AUTHORS Loftus,B., Wang,Z., Van Aken,S. and Fraser,C.
 TITLE Determination of clone end sequences from Entamoeba histolytica
 JOURNAL HM1:IMSS sheared DNA library (2001)
 COMMENT Unpublished (2001)
 Contact: Brendan J Loftus
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0208
 Fax: 301 838 3543
 Email: bjloftus@igr.org
 Clones are derived from the Entamoeba histolytica HM1:IMSS sheared

DNA library
 Seq primer: M13-Reverse
 Class: shotgun
 High quality sequence start: 18
 High quality sequence stop: 851.
 Location/Qualifiers
 1..886
 /organism="Entamoeba histolytica"
 /strain="HM1:IMSS"
 /db_xref="taxon:5759"
 /clone_lib="Entamoeba histolytica Sheared DNA"
 /note="Vector: PHOS1; Site 1: Bst I; Constructed at The Institute for Genomic Research (TIGR), Rockville, MD.
 Genomic DNA isolated from broth cultures of E. histolytica using a method described by Clark and Diamond (Clark, C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a method for isolate identification. Exp. Parasitol. 77:450.). The DNA was mechanically sheared to give a tight size distribution (~2 kb). The v + i method used for the library construction is described in detail in Smith, H.O. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B. Borell. Oxford University Press, 1999)."

FEATURES

source

BASE COUNT 348 a 105 c 176 g 257 t

ORIGIN

Alignment Scores:

Pred. No.: 1..93e-74 Length: 886
 Score: 743.50 Matches: 151
 Percent Similarity: 73.43% Conservatives: 48
 Best Local Similarity: 55.72% Mismatches: 71
 Query Match: 13.95% Indels: 2
 DB: 17 Gaps: 1

US-09-873-409-4 (1-1058) x BHI39685 (1-886)

QY 786 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 805
 DB 20 GCAAGAATACTTATGATGTTATTGATAGAAAACCACT-ATTGATTGTTATAGTGAAGA 78
 QY 806 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 825
 DB 79 GGTGAACATTTTAATGATGTTAAAGGTGAATTTAAAGACATTTGTTTCAGATAT 138
 QY 826 ProCysArgProAspValPheIleLeuArgGlyLysSerLeuSerIleGluArgGlyLys 845
 DB 139 CCAACAGACACACAAATCTGTCTTGAAGGTATTTCATTCAAGTAGACACAGGAAA 198
 QY 846 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 865
 DB 199 ACTGTTGCATTAGTAGGAGCATCAGGATGTTGTAATCAACATCAGTTCAGTTGATTGA 258
 QY 866 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 885
 DB 259 AGATTTTATGATCCAAACACATCAGAGATGATTATTAGATGGACATAATATCAAGATTG 318
 QY 886 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 905
 DB 319 AATATTCATTTCTTAAGAGTCAATTTGNAATGTAAGTAGGACAGACAGTATTATTGCT 378
 QY 906 CysSerIleAlaGluAenIleAlaTyrGlyAspAenSerArgVal- - -ValProLeuAsp 924
 DB 379 GAAAGTGTATGATAATATTAGAAGGAGGTACCTAAAGGAGTTCAGTAAGTAAGTAA 438
 QY 925 GluIleLysGluAlaAlaAenAlaAlaAenIleHisSerPheIleGluGlyLeuProGlu 944
 DB 439 CAAATTTATGCTCTGCTAAAATGGCAATGCACATCTTTATTTTCAGCAATGCCAGAA 498
 QY 945 LysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlnLysGlnArg 964
 DB 499 GGATATAACACAATGGTAGGTGATAGGTGTCACAAATTTTCAGGAGGACAAAACAAGA 558


```
Qy 965 LeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThr 984
Db 559 ATTGCTATTGCACGTGCATTGATTAGAAATCCAAAGTGTATTACTCGATGAAGCTACA 618
Qy 985 SerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThr 1004
Db 619 TCACACTTGTATTCAGAAAGTGAAGAGATTGTACAGATGCACCTTGACAAAGCAGCAAAA 678
Qy 1005 GlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIle 1024
Db 679 GGAAGACACACAATTGTAATTGCACATAGATTATCAACTATTCAAATGCAGATCAATA 738
Qy 1025 ValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsn 1044
Db 739 TGTGTTATTATGACAGCAGGAAGATTGCAGAAAGAGGACACATCAAGAGTTATTAGATTG 798
Qy 1045 ArgAspIleTyrPheLysLeuValAsnAlaGln 1055
Db 799 AAAGGATTTTATTATACACTTGCTATGCAACAA 831
```

Search completed: March 31, 2003, 13:50:31
Job time : 7305.24 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 . Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 30, 2003, 01:24:25 ; Search time 14303 Seconds
(without alignments)
2152.747 Million cell updates/sec

Title: US-09-873-409-4
Perfect score: 5330
Sequence: 1 MVISLTSKLSAYSAGAVA.....QELLNRDIYFKLVNAQSVQ 1058

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DEV=xlh
-MODEL=frame+p2n.model -DB=GenEmbl -O=/cpn2.1/USPTO spool/US09873409/runat_27032003_115418_19228/app query fasta_1.7544
-UNITs=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LOPEXT=0
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-USER=US09873409 @CN 1.1.54887 @runat_27032003_115418_19228 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sv.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.ste.*
28: em.un.*

29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	5328	100.0	3177	6	AX339030 Sequence
2	5328	100.0	3702	6	AX339031 Sequence
3	5159.5	96.8	3621	6	AX339032 Sequence
4	4985.5	93.5	3699	6	AX478104 Sequence
5	4561.5	85.6	2856	6	AX339028 Sequence
6	3456	64.8	2066	6	AX339027 Sequence
7	3137.5	58.9	4296	10	CRUGPI M60040 C.griseus p
8	3136.5	58.8	3987	10	CRUGPI1185 M9254 Chinese ham
9	3136.5	58.8	4304	10	CRUGPI1165 M9253 Chinese ham
10	3121.5	58.6	4018	5	GA9799 U78609 Ovis aries
11	3118.5	58.5	3858	4	OAU78609 AF257746 Rattus no
12	3116.5	58.5	4927	10	AF257746 Rattus no
13	3115.5	58.5	4279	6	AX105082 Sequence
14	3112.5	58.4	4279	6	AX105057 Sequence
15	3112.5	58.4	4279	6	AX105078 Sequence
16	3112.5	58.4	4279	6	AX105080 Sequence
17	3111.5	58.4	4317	4	AF045016 Canis fam
18	3111.5	58.4	4317	6	AX105059 Sequence
19	3110.5	58.4	4045	12	AF269224 Synthetic
20	3105.5	58.3	3934	4	CFA413568 AF269224 Synthetic
21	3103.5	58.2	4323	10	AF286167 Rattus no
22	3100.5	58.2	3860	6	AX322787 Sequence
23	3100.5	58.2	3860	6	AX322789 Sequence
24	3100.5	58.2	4186	6	AX108654 Sequence
25	3100.5	58.2	4195	6	AX108656 Sequence
26	3100.5	58.2	4264	6	AR051647 Sequence
27	3100.5	58.2	4264	6	AR051650 Sequence
28	3100.5	58.2	4646	6	AX336420 Sequence
29	3100.5	58.2	4646	6	AX336708 Sequence
30	3100.5	58.2	4646	6	AX391099 Sequence
31	3100.5	58.2	4646	6	I49610 Sequence 2
32	3100.5	58.2	4646	6	I49610 Sequence 2
33	3100.5	58.2	4669	6	I08557 Sequence 3
34	3100.5	58.2	6505	6	AR028671 Sequence
35	3100.5	58.2	8630	6	AX012320 Sequence
36	3100.5	58.2	8630	6	AX012321 Sequence
37	3100.5	58.2	9318	6	AR028672 Sequence
38	3097.5	58.1	4378	6	E02326 Multidrug r
39	3091	58.0	3888	10	AY082609 M24417 Mouse phosp
40	3090.5	58.0	3682	10	MUSPGLY1A AX322793 Sequence
41	3090.5	58.0	4788	6	AX322793 Sequence
42	3090.5	58.0	4924	10	MUSMDR1A M33581 Mouse p-gly
43	3089	58.0	4280	10	CRUGPII M60041 C.griseus p
44	3088.5	57.9	4356	10	MUSMDRAA M30697 Mouse multi
45	3086	57.9	4233	6	AR123273 Sequence

ALIGNMENTS

AX339030
LOCUS AX339030 3177 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 12 from Patent WO0194400.
ACCESSION AX339030
VERSION AX339030.1 GI:18129122
KEYWORDS human.
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Frank, M.H. and Sayegh, M.H.
TITLE A gene encoding a multidrug resistance human p-glycoprotein
JOURNAL homologue on chromosome 7p15-21 and uses thereof
Patent: WO 0194400-A 12 13-DEC-2001:
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES
Location/Qualifiers
source 1..3177
BASE COUNT 970 a 601 c 735 g 870 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 3177
Score: 5328.00 Matches: 1058
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.96% Indels: 0
DB: 6 Gaps: 0
US-09-873-409-4 (1-1058) x AX339030 (1-3177)
Qy 1 MetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAla 20
Db 1 ATGTCATCTCATTCAGCAGTAAAGAAATTAAGTGCCTATTCCAAAGCTGGGCTGGCA 60
Qy 21 GluGluValLeuSerSerIleAcrqThrValIleAlaPheArgAlaGlnGluLysGluLeu 40
Db 61 GAAGAGCTTGTCTCATCAATCCGAACTCATAGCCTTTAGGGCCCAAGGAAAGAACTT 120
Qy 41 GlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeu 60
Db 121 CAAGGCTTTCTTTAAATATAACAAGATATGCTTGGTTTTATTTCCTCCAGTGGCTA 180
Qy 61 LeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPhe 80
Db 181 CTAAGTTGTGTCTGTTNTTTGTAAAGGTATACACAGAATCTCAAAGATGCCAAGGATTTT 240
Qy 81 GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsn 100
Db 241 GGCAATAAAGGACTATAGCTTCAAAAGTGTCTCTGGTGTGTACTTCTTTATGAAT 300
Qy 101 GlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGly 120
Db 301 GGAACCTATGGACTTGCTTTTGGTATGAACCTCCTTGATTCTTAATGAGAACCTGGA 360
Qy 121 TyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIle 140
Db 361 TATACCATCGGAGCTGTTCTTGTCTGTTTCTTTAGTGAATCCATGACGATTTATGCATT 420
Qy 141 GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 160
Db 421 GGACGACAGTCCCTCACTTTGAAACCTTTCGCAATAGCCCCGAGGAGCTGCCTTCATATT 480
Qy 161 PheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysPro 180
Db 481 TTCCAGGTTATTGATAAGAACCCAGTATAGATAAATTTTCCACAGCTGGATATAAACCT 540
Qy 181 GluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgPro 200
Db 541 GAATCCATAGAAGAACTGTGGAATTTAAAAAATGTTCTTTCAATTTATCCATCAAGACCA 600

Qy 201 SerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu 220
Db 601 TCTATCAAGATTCTGAAGGCTCTGAATCTCAGAAATTAAGTCTGGAGAGACAGTCGCTTG 660
Qy 221 ValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAsp 240
Db 661 GTCGGTCTCAATGGCAGTGGAGAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGAT 720
Qy 241 ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis 260
Db 721 CCGGATGATGGCTTTATCATGGTGGATGAGATGACATCAGACCTTTAAATGTGGCGCAT 780
Qy 261 TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSer 280
Db 781 TATCGAGACCATATTGGAGTGGTTAGTCAAGAGCCTGTTTTGTTGGGACCAACCATCAGT 840
Qy 281 AsnAsnIleLysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArg 300
Db 841 AACATATCAAGTATGACGAGATGATGTGACTGATCAAGAGATGGAGAGCAGCAAGG 900
Qy 301 GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly 320
Db 901 GAAGCAAAATGCGPATGATTTTATCATGGAGTTTCTTAATAAATTTATACATTTGGTAGG 960
Qy 321 GluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu 340
Db 961 GAAAAAGGAGCTCAAAATGAGTGGAGGCGAGAAACACAGAGATCGCAATTGCTCGCTTA 1020
Qy 341 ValArgAsnProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspSerGluSer 360
Db 1021 GTTCGAAACCCCAAGATTCTGATTTTAGATGAGGCTACGCTCGCTCGGATTTCAGAAAGC 1080
Qy 361 LysSerAlaValGlnAlaLeuGluLysAlaSerLysGlyValArgThrThrIleValVal 380
Db 1081 AAGTCAGCTGTTCAAGCTGCACCTGGAGAGCGCAGAAAGGTGGACTCAATCGTGGTA 1140
Qy 381 AlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMet 400
Db 1141 GCACACCGACTTCTACTATTTCGAGTGCAGATTTGATTGTGACCCCTAAAGGATGGAATG 1200
Qy 401 LeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeu 420
Db 1201 CTGGCGGAGAAAGAGCAGACATGCTGAACCTAATGGCAAAACGAGGTCTATATTATTCATT 1260
Qy 421 ValMetSerGlnAspIleLysAlaAspGluGlnMetGluSerMetThrTyrSerThr 440
Db 1261 GTGATGTCAGAGATATTAAAAAGCTGTGAACAGATGGAGTCAATGACATATTCTACT 1320
Qy 441 GluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle 460
Db 1321 GAAAGAAAGACCAACTCACCTTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTCATT 1380
Qy 461 AspLysAlaGluGluSerThrGlnSerLysGluLysSerLeuProGluValSerLeuLeu 480
Db 1381 GACAAGCTCAGGAATCCACCAATCTAAAGAGATAGTCTTCTGAAAGTCTCTCTATTA 1440
Qy 481 LysIleLeuLysLeuAsnLysProGluTrpPheValValLeuGlyThrLeuAlaSer 500
Db 1441 AAAATTTTAAAGTTAAACAAGCTGAATGGCCTTTGTGGTTCTGGGACATTTGGCTTCT 1500
Qy 501 ValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMet 520
Db 1501 GTTCTAAATGGAACGTTCATCCAGTATTTTCCATCATCTTTCAAAAAATTAATAACCATG 1560
Qy 521 PheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePhe 540
Db 1561 TTTGGAAATTAATGATAAAACACCATTAAGCATGATGACAGAAATTTATTCATGATATTC 1620
Qy 541 ValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArg 560
Db 1621 GTCAATTTGGGTGTTATTTCCTTTGTCAGTTATTTTCATGCGGATTTATTTTACGGAGA 1680
Qy 561 AlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGln 580

Db 1681 GCAGGGGAAATTTAAACGATGAGATTAAGACACTTGGCCCTTCAAAGCCCATGTTATATCAG 1740
QY 581 AspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAla 600
Db 1741 GATATGGCTCGTGTGGTGAAGGAAACACACAGGAGGCTTGACAAATATTTAGCC 1800
QY 601 IleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsn 620
Db 1801 ATAGATATAGCAAAATTTCAAGGAGCAACAGGTTCAGGATTTGGCGTCTTAAACACAAAT 1860
QY 621 AlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyThrGluMetThrPhe 640
Db 1861 GCAACTAAACATGGAGCTTTTCAGTTATCAATTTCTTTATATATGATGGAGATGACATTC 1920
QY 641 LeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaIleMet 660
Db 1921 CTGATTTCTGAGTATGCTCCAGTACTTCCGCGACAGGAATGATTTGAACCCGAGCAATG 1980
QY 661 ThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGlu 680
Db 1981 ACTGGATTTGCCAACAAAGATAAGCAAGAACTTTAAGCATGCTGGAAAGATAGCAACTGAA 2040
QY 681 AlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMet 700
Db 2041 GCTTTGGAGAATATACGTACTATAGTGTCAATTAACAAAGGAAAAAGCCTTCGAGCAATG 2100
QY 701 TyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGly 720
Db 2101 TATGAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAAGAAAGCACAGATATTGGA 2160
QY 721 SerCysTyAlaPheSerHisAlaPheIleTyThrPheAlaTyAlaAlaGlyPheArgPhe 740
Db 2161 AGCTGTATGCAATTACCATGCTTATATATATTTTGCCTATGTCAGCAGGGTTTCGATTT 2220
QY 741 GlyAlaTyLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThr 760
Db 2221 CGAGCCTATTTAATCAAGCTGGACGAATGACCCAGAGGGCATGTTTCATGTTTACT 2280
QY 761 AlaIleAlaTyGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyThrSer 780
Db 2281 GCAATTTGCATATGGAGCTATGGCCATCGGAAACACCTCGTTTGGCTCTCGTAATATCC 2340
QY 781 LysAlaLysSerGlyAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAsp 800
Db 2341 AAAGCCAAATCGGGGCTGGCATCTGTTTGCCTTGTGGAAAGAAACCAATATAGAC 2400
QY 801 SerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGlu 820
Db 2401 AGCCGCAGTCNAGAGGGAAGAAAGCCAGACACATGTGAAGGGAATTTAGAGTTTCAGAA 2460
QY 821 ValSerPhePheTyProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSer 840
Db 2461 GTCTCTTTCTTCATCCATGTCGCCCAGATGTTTTCATCTCGCTGGCTTATCCCTCAGT 2520
QY 841 IleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSer 860
Db 2521 ATTGACCGAGGAAGACAGTAGGATTTTGGGGAGCAGCGGCTGTGGAAAAAGCATTTCT 2580
QY 861 ValGlnLeuLeuGlnArgLeuTyAspProValGlnGlnValLeuPheAspGlyVal 880
Db 2581 GTTCAACTTCGCAGAGACTTTATGACCCCGTCGAAGGACAAGTGTGTTGATGGTG 2640
QY 881 AspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGlu 900
Db 2641 GATGCAAAAGAAATGAAATGATACAGTGGCTCCGTTCCCAATAGCAATCGTTCCTCAAGAG 2700
QY 901 ProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyThrGlyAspAsnSerArgVal 920
Db 2701 CCTGTGCTCTTCAACTGACGATTTGCTGAGAACATCGCCTATGTTGTCAGAACACCCGTGTG 2760
QY 921 ValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGlu 940

Db 2761 GTGCATTAGATGAGATCAAAAGAGCCGAAATGCAGCAAAATATCCATCTCTTTATTGAA 2820
QY 941 GlyLeuProGluLysTyAsnThrGlnValGlyLeuLysGlyValaGlnLeuSerGlyGly 960
Db 2821 GGTCTCCCTGAGAAATACACACACAAGTTGGACTGAAAGGAGCACAGCTTTCTGGCGC 2880
QY 961 GlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeu 980
Db 2881 CAGAAACAAAGACTAGCTATTGCAAGGGCTCTTCTCCAAAAACCCAAAAATTTATTGTTG 2940
QY 981 AspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAsp 1000
Db 2941 GATGAGGCCACTTCAGCCTCGATTAATGACAGTGAAGGTGTTTCAGCATGCCCTTGAT 3000
QY 1001 LysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsn 1020
Db 3001 AAAGCCAGGACGGGAAGGACATGCTAGTGTCTCACTCACAGGCTCTCTGCAATTCAGAAC 3060
QY 1021 AlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlu 1040
Db 3061 CGAGATTTGATAGTGTGTTCTGCACAATGGAAGATAAAGGAAACAAGGAACTCATCAAGAG 3120
QY 1041 LeuLeuArgAsnArgAspIleTyThrPheLysLeuValAsnAlaGlnSerValGln 1058
Db 3121 CTCCTGAGAAATCGAGACATATATTTAAGTTAGTATGCAATGCACAGTCAGTGCAG 3174
RESULT 2
AX339031 3702 bp DNA linear PAT 09-JAN-2002
LOCUS AX339031 Sequence 13 from Patent WO0194400.
DEFINITION AX339031
ACCESSION AX339031.1 GI:18129123
VERSION AX339031.1
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Frank,M.H. and Sayegh,M.H.
TITLE A gene encoding a multidrug resistance human p-glycoprotein
homologue on chromosome 7p15-21 and uses thereof
JOURNAL Patent: WO 0194400-A 13 13-DEC-2001;
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES
source
1. 3702
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 1098 a 708 c 856 g 1039 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 3702
Score: 5328.00 Matches: 1058
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.96% Indels: 0
DB: 6 Gaps: 0
US-09-873-409-4 (1-1058) x AX339031 (1-3702)
QY 1 MetValIleSerLeuThrSerLysGluLeuSerAlaTyThrSerLysAlaGlyAlaValAla 20
Db 526 ATGGTCATCTCTATTCACCATGTAAGGAATTAAGTGCCTATTCCAAAGCTGGGCTGTGGCA 585
QY 21 GluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGlnLysGluLeu 40
Db 586 GAAGAAGTCTTGTCTCAATCCGAACAGTCATAGCCTTTAGGGCCCGAGGAAAGAACTT 645
QY 41 GlnArgSerPheLeuLeuAsnIleThrArgTyThrAlaTrpPheTyThrProGlnTrpLeu 60
Db 646 CAAAGGCTCTTCTTTTAAATATAACAAGATATGCTGTGTTTATTTCCTCCAGTGCCTA 705
QY 61 LeuSerCysValLeu***PheValArgTyThrGlnAsnLeuLysAspAlaLysAspPhe 80

|||||
706 CTAAGTGTGCTGCTGTTNTTGTAAAGGTATATACACAGAAATCTCAAGATGCAAAAGGATTTT 765
QY GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPheMetAsn 100
Db GGCAATAAAGGACTATAGCTTCNAAGTGTCTCTGGTGTGTGTACTTCTTTATGAAT 825
QY GlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAenGlyGluProGly 120
Db GGAACCTATGGACTGTCTTTTGGTATGGAACCTCTTGATCTTAATGGAGAACCTGGA 885
QY TyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIle 140
Db TATACCATCGGACTGTCTTGTCTGTTTCTTTAGTGTAAATCCATAGCAGTTATTCGATT 945
QY GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 160
Db GGAGCAGACGTCCTCTCACTTTGAAACCTTCGCAATAGCCGAGGAGCTGCTTTTCATATT 1005
QY PheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysPro 180
Db TTCAGGTTATGATAGAAACCCAGTATAGATAACTTTTCCACAGCTGGATATAACCT 1065
QY GluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgPro 200
Db GAATCCATAGAGAACTGTGGAATTTAAAAATGTTTCTTTCAATTATCCATCAAGACCA 1125
QY SerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu 220
Db TCTATCAAGATTCTGAAGGTCTGAATCTCAGAAATTAAGTCTGGAGAGACAGTCGCCCTG 1185
QY ValGlyLeuAenGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAsp 240
Db GTCGGTCTCAATGGCAGTGGAGAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGAT 1245
QY ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAenValArgHis 260
Db CCGGATGATGGCTTTATCATGGTGGATGAGAAATGACATCAGAGCTTTAAATGTGGCGCAT 1305
QY TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrIleSer 280
Db TATCGACACCATATTTGAGTGGTGGTGTAGTCAAGAGCTGTTTGTTCGGGACACCATCAGT 1365
QY AsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArg 300
Db AACATATCAAGTATGACGAGATGATGTGACTGATGAAGAGTGGAGAGCAGCAGG 1425
QY GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly 320
Db GAAGCAAAATGCGTATGATTTTATCATGGAGTTTCTTAATAAAATTAATATCATTTGGT 1485
QY GluLysGlyAlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu 340
Db GAAAAAGGAGCTCAAAATGAGTGGAGGCGCAAAACAGAGGATCGCAATTTGCTCGCCCTTA 1545
QY ValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 360
Db GTTCGAAACCCCAAGATCTGATTTTATAGATGAGGCTACGCTCGCCCTGGATTCAGAAAGC 1605
QY LysSerAlaValGlnAlaLeuGluLysAlaSerLysGlyArgThrThrIleValVal 380
Db AAGTCAGCTGTTCAAGCTGCACCTGGAGAGCGAGCAAAAGTCGGACTACAATCGTGGTA 1665
QY AlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMet 400
Db GCACACCGACTTCTACTATTCGAAGTGCAGATTGTGATTTGACCCCTAAAGGATGGAATG 1725
QY LeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeu 420
Db CTGGCGAGAAAGGAGCACATGCTGAATTAATGCAAAACGAGGCTCTATATTATTCATT 1785
QY ValMetSerGlnAspIleLysAlaAspGluGlnMetGluSerMetThrTyrSerThr 440
|||||

Db 1786 GTGATCTCAGAGATATTAAAAAGCTGTATGACACAGATGGAGTCAATGATATTTCTACT 1845
QY GluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle 460
Db GAAAGAAAGACCAACTCACTTCTCTGCACTCTGTGAAGCATCAAGTCAGACTTCATT 1905
QY AspLysAlaGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeu 480
Db GACAAGGCTGAGGAATCCACCAATCTAAAGAGATAAGTCTTCTCTGAAGTCTCTCTATTA 1965
QY LysIleLeuLysLeuAsnLysProGluTyrProPheValValLeuGlyThrLeuAlaSer 500
Db AAAATTTTAAAGTTAAACAAGCCCTGAATGCCCTTTTGTGGTCTCTGGGACATTTGGCTTCT 2025
QY ValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMet 520
Db GTTCTTAATGGAACTGTTCATCCAGTATTTTCCATCATCTTTGCAAAATATTATACCATTG 2085
QY PheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePhe 540
Db TTTGGAAATAATATGATAAAACCAACATTAAGCATGATGCAGAAATTTTATTCATGATATTC 2145
QY ValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArg 560
Db GTCATTTTGGGTGTTATTTGCTTTGTTCAGTATTTTCATGCAGGGATTTATTTACGGCAGA 2205
QY AlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGln 580
Db GCAGGGGAATTTTAACGATGAGATTAAAGACCTTGGCCCTTCAAGCCATGTTATATTCAG 2265
QY AspIleAlaTyrPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAla 600
Db GATATTTGCCCTGTTTCATGAAAGGAAACACAGCACAGGAGGCTTGACAAATATTAGCC 2325
QY IleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsn 620
Db ATAGATATATAGCAAAATTCAGGAGCAACAGGTTCCAGGATTTGGCGTCTTAAACACAAAAT 2385
QY AlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyrGluMetThrPhe 640
Db GCAACTTAACATGGGACTTTTCAGTTATCATTTCTTTATATATGATGGAGATGACATTC 2445
QY LeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMet 660
Db CTCGATTCGAGTATTGCTCCAGTACTTCCGTCGACAGGAATGATTGAAACCGCAGCAATG 2505
QY ThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGlu 680
Db ACTGGATTTGCCCAACAAAGATAGCAAGAACTTAAAGCATGCTCGAAAGATAGCAACTGAA 2565
QY AlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMet 700
Db GCTTTGGAGAAATATACGTACTATAGTGTCTATTAAAGGGGAAAAGCCCTTCGAGCAAAATG 2625
QY TyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGly 720
Db TATGAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGACACAGATTTATTGGA 2685
QY SerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPhe 740
Db AGCTGTTATGCAATTCAGCCATGCTTTTATATATTTTTCCTATGCTATGCAGCAGGTTTCGATT 2745
QY GlyValTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThr 760
Db GGAGCCTATTATTAATCAAGCTGGCAAGATGACCCAGAGGCAATGTTTCATAGTTTACT 2805
QY AlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSer 780
Db GCAATTCATATGAGCTATGGAGCTATGGCCATCGGAAACCGCTCGTTTGGCTCTCGTAATTC 2865
QY LysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAsp 800
Db AAAGCCAAATCGGGGCTCGCATCTGTTGCTTGTGGAAAGAAACCAATATAGAC 2925


```
QY 801 SerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGlu 820
Db 2926 AGCCGCAAGTCAAGAGAGGAAAGCCAGACACATGTGAAGGAAATTTAGAGTTTCAGAA 2985
QY 821 ValSerPhePheTyProCysArgProAspValPheIleLeuArgGlyLeuSer 840
Db 2986 GTCTCTTTCTTATCCATGTGCGCCAGAGTTTTCATCTCCGTGGCTTATCCCTCAGT 3045
QY 841 IleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSer 860
Db 3046 ATTGAGCGAGGAAAGACAGTAGCATTTGTGGGAGCAGCGGCTGTGGAAAGCACTTCT 3105
QY 861 ValGlnLeuLeuGlnArgLeuTyArgProValGlnGlnGlnValLeuPheAspGlyVal 880
Db 3106 GTTCAACTTCTCGAGAGACTTATGACCCCGTCAAGGACAAAGTGTGTTGATGCTGTG 3165
QY 881 AspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGlu 900
Db 3166 GATGCAAAAGAATTGAATGTACAGTGGCTCCGTTCCCAATAGCAATGCTTCTTCAAGAG 3225
QY 901 ProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyArgValAspAsnSerArgVal 920
Db 3226 CTGTGCTCTTCACTGCAGCATGTGTGAGACATCGCTATGTTGACACAGCCGTGTG 3285
QY 921 ValProLeuAspGluLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGlu 940
Db 3286 GTGCCATTAGATGAGATCAAGAAAGCGCAATGCGCAAAATATCCATTCTTTTATTGAA 3345
QY 941 GlyLeuProGluLysThrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGly 960
Db 3346 GTCTCCCTCGAGAAATACAAACACACAAGTTGGACTGAAAGGAGCACAGCTTTCTGCGGC 3405
QY 961 GlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeu 980
Db 3406 CAGAAACAAGACTAGCTATTGCAAGGCTCTTCTCCAAAACCCAAATTTATTGTTG 3465
QY 981 AspGluAlaThrSerAlaLeuAspAsnSerGluLysValValGlnHisAlaLeuAsp 1000
Db 3466 GATGAGGCCACTTCAGCCCTCGATATATGACAGTGAAGAGTGGTTCCAGCATGCCCTTGAT 3525
QY 1001 LysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsn 1020
Db 3526 AAAGCCAGGACGGGAGGACATGCTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAAC 3585
QY 1021 AlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlu 1040
Db 3586 GCAGATTGATAGTGTCTGCACAAATGGAAGATGAAGAAACAGGAACCTCATCAAGAG 3645
QY 1041 LeuLeuArgAsnArgAspIleTyPheLysLeuValAsnAlaGlnSerValGln 1058
Db 3646 CTCCTGAGAAATCGAGACATATATTTTAAGTTAGTGAATGCACAGTCAGTCGACG 3699

RESULT 3
LOCUS AX339032 3621 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 14 from Patent WO0194400.
ACCESSION AX339032
VERSION AX339032.1 GI:18129124
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Frank, M.H. and Sayegh, M.H.
TITLE A gene encoding a multidrug resistance human p-glycoprotein
homologue on chromosome 7p15-21 and uses thereof
JOURNAL Patent: WO 0194400-A 14 13-DEC-2001;
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES
source Location/Qualifiers
1..3621
/organism="Homo sapiens"
```

```
BASE COUNT 1081 a 696 c 842 g 1002 t
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 3621
Score: 5159.50 Matches: 1031
Percent Similarity: 97.45% Conservative: 0
Best Local Similarity: 97.45% Mismatches: 0
Query Match: 96.80% Indels: 27
DB: 6 Gaps: 1
US-09-873-409-4 (1-1058) x AX339032 (1-3621)
QY 1 MetValIleSerLeuThrSerLysGluLeuSerAlaTySerLysAlaGlyAlaValAla 20
Db 526 ATGGTCACTCTCATTGACCAGTAAGGAATTAAGTGCCTATTCCAAAGCTGGGCTGTGGCA 585
QY 21 GluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGlnLysGluLeu 40
Db 586 GAAGAAGTCTTTGTTCATCAATCCGAACAGTCATAGCTTTTAGGGCCCGAGGAAGAAGACTT 645
QY 41 GlnArgSerPheLeuLeuAsnIleThrArgTyAlaTrpPheTyPheProGlnTrpLeu 60
Db 646 CAA----- 648
QY 61 LeuSerCysValLeu***PheValArgTyThrGlnAsnLeuLysAspAlaLysPhe 80
Db 649 -----AGGTATACACAGAATCTCAAAGATGCAAAAGGATTTT 684
QY 81 GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyPhePheMetAsn 100
Db 685 GGCATAAAAGAGACTATAGCTTCAAAGTGTCTCTTGGTGTGTGTACTTCTTTATGAAT 744
QY 101 GlyThrTyGlyLeuAlaPheTrpTyGlyThrSerLeuIleLeuAsnGlyGluProGly 120
Db 745 GGAACCTATGGAAGTCTTTTGGTATGGAACCTCTTGATTTCTTAATGAGAGACCTGGA 804
QY 121 TyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyCysIle 140
Db 805 TATACCATCGGGAGCTTCTTGTCTTTCTTTAGTGAATCCATAGCAGTTATTGCAAT 864
QY 141 GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 160
Db 865 CGAGCAGCAGTCCCTCAGTTCGAAACCTTCGCAATAGCCGAGGAGCTGCTTCATATT 924
QY 161 PheGlnValIleAspLysProSerIleAspAsnPheSerThrAlaGlyTyLysPro 180
Db 925 TTCCAGGTTATTGATAAGAAACCCAGTATAGATAACTTTTCCACAGCTGGATATAAACCT 984
QY 181 GluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyProSerArgPro 200
Db 985 GAATCCATAGAAGGAACCTGGGAATTTAAATGTTCTTTCAATATCCATCAAGACCA 1044
QY 201 SerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu 220
Db 1045 TCTATCAAGATTTCTGAAAGGCTGTAATCTCAGAAATTAAGTCTGAGAGACAGCTGCGCTTG 1104
QY 221 ValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyArg 240
Db 1105 GTCCGTCTCAATGGCAGTGGGAGAGTAGCGGTAGTCCAGCTTCTCGAGAGGTTATATGAT 1164
QY 241 ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis 260
Db 1165 CCGGATGATGGCTTTTATCATGTGTGATGAGATGACATCAGAGCTTTAAATGTGGGCAT 1224
QY 261 TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSer 280
Db 1225 TATCGAGACCAATATTGGAGTGGTGTAGTCAAGAGCCTGTTTGTTCGGGACCACCATCACT 1284
QY 281 AsnAsnIleLysTyGlyArgAspValThrAspGluGluMetGluArgAlaAlaArg 300
Db 1285 AACAAATATCAAGTATGGAGCAGATGATGTGCTGATGAAGAGATGGAGAGAGCAGCAAG 1344
```


QY 301 GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly 320
DB 1345 GAAGCAAAATCGGTATGATTTTATCATGGAGTTTCTTAATAATTTAATACATGGTAGGG 1404
QY 321 GluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu 340
DB 1405 GAAAAAGGAGCTCAAAATGAGTGGAGGCGAGAAACAGAGGATCGCAATTGCTCGTGCCTTA 1464
QY 341 ValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 360
DB 1465 GTTCGAAACCCCAAGATTTCTGATTTTAGATGAGGCTACGCTGCGCTGGATTCAGAAAGC 1524
QY 361 LysSerAlaValGlnAlaLeuGluLysAlaSerLysGlyArgThrThrIleValVal 380
DB 1525 AAGTCAGCTGTTCAAGCTGCATCGGAGAGGCGAGCAAGGTGGACTACATCGTGGTA 1584
QY 381 AlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMet 400
DB 1585 GCACACCGACTTTCTACTATTTCGAAGTGCAGATTTTGATTGTGACCTAAAGGATGGAATG 1644
QY 401 LeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeu 420
DB 1645 CTGGCGGAGAAAGGAGCACATGCTGAACATAATGCGCAAAACGAGGTCTATATTATTACATT 1704
QY 421 ValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThr 440
DB 1705 GTGATGTACAGATATTAAAGAGCTGATGAACAGATGGAGTCAATGACATATTTCTACT 1764
QY 441 GluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle 460
DB 1765 GAAAGAAAGACCAACTCACITTCCTGCACTCTGTGAAGAGCATCAAGTCAGACTTCAATT 1824
QY 461 AspLysAlaGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeu 480
DB 1825 GACAAGGCTCAGGAATCCACCAATCTAAAGAGATAAGTCTTCTGAAAGTCTCTCTATTATTA 1884
QY 481 LysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSer 500
DB 1885 AAAATTTTAAAGTTAAACAAGCTGAATGGCTTTTGTGGTTCTGGGACATTTGGCTTCT 1944
QY 501 ValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMet 520
DB 1945 GTTCTAAATGGAAGTGTTCATCCAGTATTTTCCATCATCTTTGCAAAATTTATAACCATG 2004
QY 521 PheClyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePhe 540
DB 2005 TTTGGAAATAATGATAAAACCCACATTAAGCATGATGACAGAAATTTTATTCATGATATTC 2064
QY 541 ValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArg 560
DB 2065 GTCAATTTGGGTGTTATTTGCTTTGTGTCAGTTATTTTCATGAGGATTTATTTTACGGCAGA 2124
QY 561 AlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGln 580
DB 2125 GCAGGGGAAATTTTAAAGATGAGATTAAGACACTTTGGCTTTCAAGCCATGTTATATCAG 2184
QY 581 AspileAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAla 600
DB 2185 GATATTCCTCGTGGTTTGTATGAAGAGGAAAAACAGCAGAGGCTTGCAACAAATATTAGCC 2244
QY 601 IleAspileAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsn 620
DB 2245 ATGATATAGCACAAATTAAGGAGCAACAGGTTCCAGGATTTGGCTTCTTAAACAAATAT 2304
QY 621 AlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPhe 640
DB 2305 GCAACTAACATGGGACTTTCAGTTTATCATTTCTTTATATATGATGGATGGGAGATGACATTC 2364
QY 641 LeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMet 660
DB 2365 CTGATTTCTGATATTGCTCCAGTACTTTGCCGTGACAGGAATGATTTGAACCGCAGCAATG 2424

QY 661 ThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGlu 680
DB 2425 ACTGATTGTCGCAACAAAGATAAGCAAGAACTTAAGCATGCTGGAAGATAGCAACTGAA 2484
QY 681 AlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMet 700
DB 2485 GCTTTGGAGAAATATACGTACTATAGTGTCTTAAACAAGGGAAAAAGCCTTCGAGCAATG 2544
QY 701 TyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGly 720
DB 2545 TATGAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCACAGATTATGGA 2604
QY 721 SerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPhe 740
DB 2605 AGCTGTTATGCAATTCAGCCATGCTTTATATATTTTGCCTATGCAGCAGGGTTTCGATTT 2664
QY 741 GlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThr 760
DB 2665 GGAGGCTATTTAAATCAAGCTGGAGCAATGACCCAGAGGGCATGTTTCATAGTTTITACT 2724
QY 761 AlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSer 780
DB 2725 GCATTTGCATATGGAGCTATGGCCATCGGAAAAACGCTGTTTGGCTCTCTGAATATTC 2784
QY 781 LysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAsp 800
DB 2785 AAGCCAAATCGGGGCTGCGCATCTGTTTGCCTTGTGGAAAAAGAAACCAATATAGAC 2844
QY 801 SerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGlu 820
DB 2845 AGCCCGCAGTCAAGAAAGGAAAAAGCCAGACACATGTGAAGGGAATTTAGAGTTTCAGAA 2904
QY 821 ValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSer 840
DB 2905 GTCCTTTCTTCTATCCATGTCGCCAGATGTTTTCATCTCCGTGGCTTATCCCTCAGT 2964
QY 841 IleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSer 860
DB 2965 ATTGACGAGGAAAGACACAGTACATTTTGTGGGAGCAGCGGCTGTGGGAAAGACATCTCT 3024
QY 861 ValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlnValLeuPheAspGlyVal 880
DB 3025 GTTCAACTTCTCAGAGACTTTATGACCCCGTGAAGGACAAAGTGTCTGTTGATGTGTG 3084
QY 881 AspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGlu 900
DB 3085 GATGCAAAAGAAATGATGATGACGGCTCGCTCCGTTCCCAATAGCAATCGTCTCTCAAGAG 3144
QY 901 ProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgVal 920
DB 3145 CCTGTGCTCTTCAACTGCAGCATTTGCTGAGAACATCGCTATGCTGACAAACGCCGTGTG 3204
QY 921 ValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGlu 940
DB 3205 GTGCCATTAGATGAGATCAAGAGCGCAATGCGCAAAATGCGCAAAATATCCATCTTITATGAA 3264
QY 941 GlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyValaGlnLeuSerGlyGly 960
DB 3265 GGTCTCCCTCAGAAATACAAACACAAAGTTGGACTGAAAGGAGCACAGCTTCTCGCGCGC 3324
QY 961 GlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeu 980
DB 3325 CAGAAACAAAGACTAGCTATTCAGGGCTCTTCTCCAAAAACCCAAAAATTTTATTTGTTG 3384
QY 981 AspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAsp 1000
DB 3385 GATGAGGCCACTTCAGCCCTCGAATAATGACGTGAGAGGTGTTTCAGCATGCCCTTGAT 3444
QY 1001 LysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsn 1020
DB 3445 AAGCCAGGACGCGAAGGACATGCTTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAAC 3504
QY 1021 AlaAspLeuIleValValLeuHisAsnGlyLysIleGluGlnGlyThrHisGlnGlu 1040


```
Db 3505 GCAGATTGATAGTGGTCTGCACATAGGAAGATAAAGAAACAAGAACTCATCAAGAG 3564
|||||
Qy 1041 LeuLeuArgAsnArgAspIleTyrPheLeuValAsnAlaGlnSerValGln 1058
|||||
Db 3565 CTCTCGAGAAATCGAGACATATATTTTAAGTTAGTGAATGCACAGTCAGTCAG 3618
|||||

RESULT 4
AX478104
LOCUS AX478104
DEFINITION Sequence 34 from Patent WO240541.
ACCESSION AX478104
VERSION AX478104.1 GI:22217064
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1
AUTHORS Tang, Y.T., Yue, H., Nguyen, D.B., Hafalia, A.J., Elliott, V.S., Lu, Y.,
Walia, N.K., Yao, M.G., Baughn, M.R., Gandhi, A.R., Ding, L.,
Sanjanwala, M., Ramkumar, J., Arvizu, C., Gietzen, K.J., Lal, P.G.,
Azinzai, Y., Khan, F.A., Thangavelu, K., Thornton, M., Lu, D.A.,
Tribouley, C.M., Warren, B.A., Ison, C.H., Das, D., Raumann, B.E.,
Policky, J.D. and Kearney, L.
TITLE Transporters and ion channels
JOURNAL Patent: WO 0240541-A 34 23-MAY-2002;
Incyte Genomics, Inc. (US)
FEATURES
source 1..3699
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="Incyte ID No: 7472030CB1"
BASE COUNT 1116 a 707 c 860 g 1016 t
ORIGIN

Alignment Scores:
Pred. No.: 26-317 Length: 3699
Score: 4985.50 Matches: 1002
Percent Similarity: 94.80% Conservative: 1
Best Local Similarity: 94.71% Mismatches: 2
Query Match: 93.54% Indels: 53
DB: 6 Gaps: 2

US-09-873-409-4 (1-1058) x AX478104 (1-3699)
Qy 1 MetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAla 20
|||||
Db 682 ATGGTCATCTCATTCATGACCAAGTAAGAAATTAAGTGCCTATTTCCTCAAGCTGGGCTGTGGCA 741
|||||
Qy 21 GluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeu 40
|||||
Db 742 GAAGAAGTCTTGTCATCAATCCGACAGTCATAGCCCTTAGGGCCCGGAGGAAAGAACTT 801
|||||
Qy 41 GlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeu 60
|||||
Db 802 CAA----- 804
Qy 61 LeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPhe 80
|||||
Db 805 -----AGGTATACACAGAAATCTCAAGATGCAAAAGGATTTT 840
Qy 81 GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsn 100
|||||
Db 841 GGCATATAAAGGACTATAGCTTCAAAAGTGCTCTCTGGTGCTGTGACTTCTTTATGAAT 900
|||||
Qy 101 GlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGly 120
|||||
Db 901 GGAACCTATGGACTTGCTTTTGGTATGGAACCTCCTTGATTCTTAATGAGAACTGGA 960
|||||
Qy 121 TyrThrIleGlyThrValLeuAlaValPheSerValIleHisSerSerTyrCysIle 140
|||||
Db 961 TATACCATCGGACTGTCTTCTGTCTTTCTTTAGTGTATTCATTCATAGTATTGCAATT 1020
|||||
```

```
Qy 141 GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 160
|||||
Db 1021 GGAGCAGCAGTCCCTCCTCCTTTCGAAACCTTCGCAATAGCCGAGAGCTGCCCTTCATATT 1080
|||||
Qy 161 PheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysPro 180
|||||
Db 1081 TTCCAGGTTATTGATAAGAAACCCAGTATAGGTAACTTTTCCACAGCTGGATATAAACCT 1140
|||||
Qy 181 GluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgPro 200
|||||
Db 1141 GAATCCATAGAAGAACTGTGGAATTTAAATAATGTTCTTTCAATTATCCATCAAGGCCA 1200
|||||
Qy 201 SerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu 220
|||||
Db 1201 TCTATCAAGATTCTGAAGGCTCTGAATCTCGGAATTAAGTCTGGAGAGACAGTCGCGCTTG 1260
|||||
Qy 221 ValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuGlnArgLeuTyrAsp 240
|||||
Db 1261 GTCCGTCTCAATGGCAGTGGGAGAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGAT 1320
|||||
Qy 241 ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis 260
|||||
Db 1321 CCGGATGATGGCTTTATCATGTGTGATGATGATGATGATGATGATGATGATGATGAT 1380
|||||
Qy 261 TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSer 280
|||||
Db 1381 TATCGAGACCATATTGGAGTGTAGTCAAGAGCCTGTTTGTTCGGGACCAACCATCAGT 1440
|||||
Qy 281 AsnAsnIleLysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArg 300
|||||
Db 1441 AACAAATATCAAGTATGACGAGATGATGTGACTGATGAAGAGATGGAGAGCAGCAAG 1500
|||||
Qy 301 GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly 320
|||||
Db 1501 GAAGCAATCGGTATGATTTTATCATGGATTTTCTTAATAATTTAATACATTTGGTAGG 1560
|||||
Qy 321 GluLysGlyAlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu 340
|||||
Db 1561 GAAAGAGAGCTCAAATGATGGAGGCGAGAAACAGAGATCGCAATTCCTCGTGCCTTA 1620
|||||
Qy 341 ValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 360
|||||
Db 1621 GTTCGAAACCCCAAGATTCTGATTTTAGTAGGCTACGCTCTGCGCTGGATTTCAGAAAGC 1680
|||||
Qy 361 LysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValVal 380
|||||
Db 1681 AAGTCAGCTGTTCAAGCTGCACTGGAGAGGCGAGCAAGGTCGGACTACAACTCGTGGTA 1740
|||||
Qy 381 AlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMet 400
|||||
Db 1741 GCACACCGACTTCTTACTATTTCGAAGTCAGATTTGATTGTGACCCCTAAAGGATGGAATG 1800
|||||
Qy 401 LeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeu 420
|||||
Db 1801 CTGGCGGAGAAAGGCGCATGCTGAACTTAATGCGAAACGAGGCTCTATATTATTCACTT 1860
|||||
Qy 421 ValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThr 440
|||||
Db 1861 GTGATGTCACAGGATATAAAGCTGATGAACAGATGGAGTCAATGACATATTCTACT 1920
|||||
Qy 441 GluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle 460
|||||
Db 1921 GAAAGAAAGCACTCACTTCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTCATT 1980
|||||
Qy 461 AspLysAlaGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeu 480
|||||
Db 1981 GACAAGGCTGAGGAATCCACCAATCTAAAGAGATAAGTCTTCTCCTGAAGTCTCTCTATTA 2040
|||||
Qy 481 LysIleLeuLysLeuAsnLysProGluTrpPheValValLeuGlyThrLeuAlaSer 500
|||||
Db 2041 AAAATTTTAAAGTTAAACAAGCCTGAATGGCCCTTTTGTGGTGTCTGGGGACATTTGGCTTCT 2100
|||||
```



```
QY 501 ValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMet 520
Db 2101 GTTCTAAATGGAACGTGTTCACTCCAGTATTTTCCATCATCTTTGCAAAAATTAATAACCATG 2160
QY 521 PheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePhe 540
Db 2161 TTTTGGAAATTAATGATTAACACACATTAAGCATGATGCAGAAAATTTATTCCATGATATTC 2220
QY 541 ValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArg 560
Db 2221 GTCAATTTGGGTGTTATTTGCTTTGTGAGTATTTTCAAGT----- 2259
QY 561 AlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGln 580
Db 2260 -----CAG 2262
QY 581 AspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAla 600
Db 2263 GATATTTGCTTGGTTTGTATGAAAGGAAACAGCACAGAGGCTTGACAAATATTAGCC 2322
QY 601 IleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsn 620
Db 2323 ATAGATATAGCACAAATTCAGAGGACCAACAGGTTCCAGGATTCGCGTCTTTAAACACAAAT 2382
QY 621 AlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyrGluMetThrPhe 640
Db 2383 GCAACTAACATGGGACTTTTCAGTTATCATTTCTTTATATATGATGGAGATGACATTC 2442
QY 641 LeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMet 660
Db 2443 CTGATTTCTGAGTATTCCTCAGTACTTTCGCGTGCACAGGAATGATTGAAACCGCAGCAATG 2502
QY 661 ThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGlu 680
Db 2503 ACTGGATTTGCCAACAAAGATAAGCAGAACTTAAGCATGCTGGAAGATAGCAACTGAA 2562
QY 681 AlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMet 700
Db 2563 GCTTTGGAGATATACGTACTATAGTGTCTATTAAACAGGGGAAAAGCCTTCGAGCAATG 2622
QY 701 TyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGly 720
Db 2623 TATGAAGAGATGCTTTCAGACTCAACACAGAAATACCTCGAAGAAAGCACAGATTTATGGA 2682
QY 721 SerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPhe 740
Db 2683 AGCTGTTATGCATTCAGCCATGCTCTTATATATTTTCCCTATGCGGAGGGTTTCGATTT 2742
QY 741 GlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThr 760
Db 2743 GGAGCCTATTAAATTCAGCTGGACGATGACCCACAGGGGCGATGTTTCATAGTTTACT 2802
QY 761 AlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSer 780
Db 2803 GCAATTCATATGAGCTATGGCCATCGGAGAAACGCTGTTTGGCTCTCGATATATCC 2862
QY 781 LysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAsp 800
Db 2863 AAACCCAAATCGGGGCTGCGCATCTGTTGCTTGTGGAAAAGAACCAATATATAGC 2922
QY 801 SerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGlu 820
Db 2923 AGCCGCGAGTCAAGAGGGGAAAACCCAGACACATGTGAGGGGATTTAGAGTTTCGAGAA 2982
QY 821 ValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSer 840
Db 2983 GTCCTTTCTTCTATCCATGTCGCCAGATGTTTTCATCTCTCCGTCGCTTATCCCTCAGT 3042
QY 841 IleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSer 860
Db 3043 ATTGAGCGAGGAAAAGACATGATGATTTTGGGGAGCAGCGCTGTGGAAAAGCACTTCT 3102
QY 861 ValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyVal 880
```

```
Db 3103 GTTCAACTTCTCAGAGACTTTATGACCCCGTGCAGGCAAGTCTGTTTATGCTGTG 3162
QY 881 AspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGlu 900
Db 3163 GATGCAAAAGAAATTAATGATGACATGGCTCGGTTCCCAANTAGCAATCGTTCTCAAGAG 3222
QY 901 ProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgVal 920
Db 3223 CCTGTGCTCTTCAACTGCAGCATTTGCTGAGAACATCGCCTATATGTTGACAAACCGCTGTG 3282
QY 921 ValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGlu 940
Db 3283 GTGCCATTAGATGAGATCAAGAACGCCAAATGCAGCAAAATATCCATTCTTTATTGAA 3342
QY 941 GlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyValaGlnLeuSerGlyGly 960
Db 3343 GGTCTCCCTCGAGAAATACCAACACACAAATTTGGACTGAAAGGAGCACAGCTTTCTGCGCGC 3402
QY 961 GlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeu 980
Db 3403 CAGAAACAAAGACTAGCTATTGCAAGGGCTCTTCTCCAAAAACCCAAAAATTTATTGTTG 3462
QY 981 AspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAsp 1000
Db 3463 GATGAGCCACTTCAGCCCTCGATAATGACAGTGAAGAGTGGTTTCAGCATGSCCTTGAT 3522
QY 1001 LysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsn 1020
Db 3523 AAAGCCAGACCGGAAAGGACATGCCCTAGTGGTCACTTCACAGGCTCTCTCAATTCAGAAC 3582
QY 1021 AlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlu 1040
Db 3583 GCAGATTTCATAGTGTCTTCGACATGGAAGATAAAGAACAGAACTCATCAAGAG 3642
QY 1041 LeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 1058
Db 3643 CTCCTGAGAAATCGACATATATTTTAAAGTTAGTGAATGCACAGTCAGTCAGTCAG 3696

RESULT 5
AX339028 2856 bp DNA linear PAT 09-JAN-2002
LOCUS Sequence 10 from Patent WO0194400.
DEFINITION AX339028
ACCESSION AX339028
VERSION AX339028.1 GI:18129120
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1
AUTHORS Frank, M.H. and Sayegh, M.H.
TITLE A gene encoding a multidrug resistance human p-glycoprotein
homologue on chromosome 7p15-21 and uses thereof
JOURNAL Patent: WO 0194400-A 10 13-DEC-2001;
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES
source
1..2856
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 888 a 540 c 652 g 776 t
ORIGIN

Alignment Scores:
Pred. No.: 7,66e-290 Length: 2856
Score: 4561.50 Matches: 912
Percent Similarity: 97.96% Conservative: 0
Best Local Similarity: 97.96% Mismatches: 1
Query Match: 85.58% Indels: 18
DB: 6 Gaps: 1

US-09-873-409-4 (1-1058) x AX339028 (1-2856)
```


QY 129 ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHis-PheGI 148
DB 114 GTTTCCTTTAGTGTAATCCATAGCAGTATTATGCAATGGAGCAGCAGCTCCCTCATTTAT -- 171
QY 148 UthrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPr 168
DB 172 -----GATAAGAAACC 182
QY 168 oSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGI 188
DB 183 CAGTATAGATAACTTTCCACAGCTGGATATAAACCTGAATCCATAGAGGAACCTGTGA 242
QY 188 upheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLysGlyLe 208
DB 243 ATTTAAATAATGTTCTTTCAATTTATCCATCAAGACCATCTATCAAGATTCTGAAGGTCT 302
QY 208 uAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLy 228
DB 303 GAATCTCAGAAATTAAGTCTGGAGAGACGTGGCTTTGGTGGTCTCAATGGCAGTGGAA 362
QY 228 sSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVa 248
DB 363 GAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCGGATGATGGCTTTATCATGT 422
QY 248 lAspGluAsnAspIleArgAlaIleuAsnValArgHisTyrArgAspHisIleGlyValVa 268
DB 423 GGATGAGAAATGACATCAGAGCTTTAAATGTGGCGCATTTATCAGAGCCATATTGGAGTGT 482
QY 268 lSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAs 288
DB 483 TAGTCAAGAGCCTTTTGTTCGGAGCACCATCATGAACAATATCAAGTATGGAGAGA 542
QY 288 pAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheI 308
DB 543 TGATGTGACTGATGAAGAGATCGAGAGACAGCAAGGGAAGCAAAATGCGTATGATTTAT 602
QY 308 eMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGI 328
DB 603 CATGGAGTTTCCTAATAAATTTAATACATTTGGTAGGGGAAAAGGAGCTCAAAATGAGTGG 662
QY 328 vGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuI 348
DB 663 AGGGCAAGAACAGAGATCGCAATTCCTGTCCTTAGTTCGAAACCCCAAGATTCCTGAT 722
QY 348 eLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLe 368
DB 723 TTTAGTAGAGGTACCTCTGCCCTGGATTTCAGAAACCAAGTCAGCTGTTCAAGCTGCAC 782
QY 368 uGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIleAr 388
DB 783 GGAGAAAGCGCAGCAAAAGGTCCGACTACAATCGTGTGTAGCACACCGACTTTCTACTATTGG 842
QY 388 gSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAl 408
DB 843 AAGTGCAGATTTGATTTGTGACCCCTAAAGGATGGAATGCTGGCGGAGAAAGGACATGC 902
QY 408 aGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIleLysLy 428
DB 903 TGAACCTAATGGCAAAACAGGCTCTATATTATTCATCTTGATGTCACAGATATTAAAAA 962
QY 428 sAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuPr 448
DB 963 AGCTGATGAACAGATGGAGTCAATGACATATTCTACTGAAAGAACCAACTCACTTCC 1022
QY 448 oLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThrGI 468
DB 1023 TCTGCACTCTGTGAAGAGCATCAAGTCAGACTTCATTTGACAGGCTGGAGGAATCCACCCA 1082
QY 468 nSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysPr 488
DB 1083 ATCTAAAGAGATTAAGTCTCTCTGGAAGTCTCTCTATTAAAAATTTTAAAGTTAAACAAGCC 1142
QY 488 oGluTrpProPheValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPr 508

DB 1143 TGAATGGCCCTTTTGGTCTTGGGACATTTGGCTTCTGTCTTAAATGGAACTGTTCATCC 1202
QY 508 oValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrTh 528
DB 1203 AGTATTTTCCATCATCTTTGCAAAAATTTATACCATGTTTGGAAATATGATATAAACCCAC 1262
QY 528 rLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPh 548
DB 1263 ATTTAAAGCATGATGCAAAATTTATTCATGATATTTCGTCAATTTGGGTGTTATTGCTT 1322
QY 548 eValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetAr 568
DB 1323 TGTCACTATTATTTTCATGCGAGGATTTTACGGCAGAGCAGGGGAAATTTTAAACGATGAG 1382
QY 568 gLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGlyLy 588
DB 1383 ATTTAGACACTTGGCTTCAAGCCATGTTATATCAGGATTTGCTGGTTTGATGAAAA 1442
QY 588 sGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGI 608
DB 1443 CGAAACACAGCACAGGAGCTTGACAACAATATTAGCCATAGATATAGCACAAATTCAGG 1502
QY 608 yAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerVa 628
DB 1503 AGCAACAGGTTCCAGGATTTGGCTCTTAACACAAAATGCAACTAACATGGGACTTTCACT 1562
QY 628 lIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProVa 648
DB 1563 TATCATTTCTTTATATATGATGGAGATGACATTTCTGATTTCTGATTTGCTCCAGT 1622
QY 648 lLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLy 668
DB 1623 ACTTCCCGTGACAGGAATGATTGAAACCGCAGCAATGACTGGATTGTCACAAAGATAA 1682
QY 668 sGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIl 688
DB 1683 GCAGAACACTTAAGCATGCTGGAAAGATAGCAACTGAAGCTTTGGAGAAATATACGTACTAT 1742
QY 688 eValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGI 708
DB 1743 AGTGTCATTTAACAGGGAAGAAAGCTTCGAGCAATGATGATGAAGAGATGCTTCAGACTCA 1802
QY 708 nHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAl 728
DB 1803 ACACAGAAATACCTCGAAGAAAGCACAGATTTATGGAAGCTGTATGCAATTCAGCCATGC 1862
QY 728 aPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGI 748
DB 1863 CTTTATATATTTTGGCTATGTCAGCAGGGTTTCGATTTGGAGCTTATTTAATTCAGCTGG 1922
QY 748 yArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAl 768
DB 1923 ACGAATGACCCAGAGGCGATGTTTCATAGTTTTTACTGCATTTGCAATTCATGAGCTATGCG 1982
QY 768 alGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHi 788
DB 1983 CATCGAAAAACGCTCGTTTTGGCTCTGAAATATTCCAAAGCCAAATCGGGGCTGCGCA 2042
QY 788 sLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLy 808
DB 2043 TCTGTTTGGCTTTGTTGGAAAAAGAAACCAAAATATAGACGCCGCGCATGCAAGAGGAAAAA 2102
QY 808 sProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysAr 828
DB 2103 GCCAGACATGTTGAAGGAATTTTAGAGTTTCGAGAAGTCTCTTCTTCTATTCATGCTCG 2162
QY 828 gProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAl 848
DB 2163 CCCAGATGTTTTTCATCCTCCGTTGCTTATCCCTCAGTATTGAGCGAGGAAAGACAGTAGC 2222
QY 848 aPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTy 868

Db 918 ----- 918
Qy 64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspPheGlyIleLys 83
Db 919 -----AGGTCAACAACAACTTTAGAGAGCCAAAAGACTTGGGATAAG 963
Qy 84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPheMetAsnGlyThrTyr 103
Db 964 AAGCGATCACAGCCACATTTCCATGGGCGCAGCTTCTTATCTATGCATCATAT 1023
Qy 104 GlyLeuAlaPheThrTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyThrIle 123
Db 1024 GCTCTGGCTTCTGGTATGGAACTCTCTGGTATCTCAAAAGAA-----TATCTATT 1077
Qy 124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
Db 1078 GGCAAGTGCTCACTGCTCTTCTTGGTGTATTAATGGGCGATTCAGTATTTGGACGGCA 1137
Qy 144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163
Db 1138 TCTCCAATATCGAAGCTTTGGCAATGCAAGAGAGAGAGAGAGCTATGAATCTTCAACATA 1197
Qy 164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
Db 1198 ATTGATAATAAGCCAGTATTTGACAGCTTCTCAAGAAGTGGTACAAACCCAGACACAT 1257
Qy 184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
Db 1258 AAGGAAATTTGGAAATTCAAAATATTCACCTTCAGTTACCCATCTCGAAAGACGTTTCAG 1317
Qy 204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
Db 1318 ATCTTGAGGGCTCACTGAGAGGTGAGAGCGGACGACAGAGTGGCCCTGGTTGGCAAC 1377
Qy 224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243
Db 1378 AGTGGCTGTGGGAAAGCACCACCTGCTCCAGCTGTGCAGAGGCTCTACGACCCACAGAG 1437
Qy 244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
Db 1438 GCGGTGTGCTAGTATCGACGGTTCAGGACATCAGGACCAATCAATGTGAGGTATCTCGGGAA 1497
Qy 264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283
Db 1498 ATCATTTGGGTCGTGAGTCAGAACCTGTGTGTTGCCACCACTATCGCTGAAACATT 1557
Qy 284 LysTyrGlyArgAspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsn 303
Db 1558 CGCTATGGCCGAGAAATGTCACTATGACGAGATCGAGAAAGCTGTCAAGGAAGCCAAAT 1617
Qy 304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323
Db 1618 GCCTATGACTTTCATCATGAATGCCCCCAATTTGACACTGTGTGGTGAGAGAGGA 1677
Qy 324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343
Db 1678 GCACAGCTGAGTGGAGGACAGAAACAGAGATCGCCATTCGTCGCTCGCTCGGCAAC 1737
Qy 344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
Db 1738 CCCAAGATCTTTGTTGGATGAGGCGACATCAGCTTGGACACAGAAAGTGAAGCCGCTG 1797
Qy 364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383
Db 1798 GTTCAGGCTGCTCTGGATTAAGCTAGAGAGCGCGGACTACCATTTGTGTAGCCACCGC 1857
Qy 384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403
Db 1858 TTGCTACAGTTTCGAAATCTCGACATCATAGCTGGTTGATGGTGTGCTCATTTGGAG 1917
Qy 404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423
Db 423 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423

Db 1918 CRAAGAAATCATGAAGAGCTCATGAGAGAAAGGGCAATTTACTTCAAACTTGTCATGACA 1977
Qy 424 Gln-----AspIleLysLysAlaAspGluGlnMet 433
Db 1978 CAGACAGCAGGAAATGAAATTTAGAAATGAAGTTGGTCAGTCTAAAATGAAAT 2037
Qy 434 GluSerMetThrTyrSerThrGluArgLysThrAsnSerLeu-----ProLeuHisSer 451
Db 2038 GATAATTTAGACATGCTTCAAAAGATTCAGCATCCAGCTTAATCAGAAAGAGATCAACT 2097
Qy 452 ValLysSerIleLysSerAspPheIleAspLysAlaGluLysSerThr---GlnSerLys 470
Db 2098 CGCAGAAGTATCGTGGACCATGACCAAGCAGAGAGCTTAGTACCAAGAGCCCTG 2157
Qy 471 GluIleSerLeuProGluValSerLeuLysIleLeuLysLeuAsnLysProGluTyr 490
Db 2158 GATGAAGATGACTCTCAATTTCTTTGGCGTATCTCTAAAGTTGAATTCATCTGAATGG 2217
Qy 491 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 510
Db 2218 CCATATTTTGTGGTGTATATTCGTGCCATAGTAATGGAGCCTTGCAACACGACATTC 2277
Qy 511 SerIlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAspLysThrLeu 529
Db 2278 TCAATATATTTCTAAGGTTGTAGGGTTTTTCAAGAATAACTGATGATGAACCAAA 2337
Qy 530 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 549
Db 2338 CGACATGATAGCAACTGTTCCTTATTTCTGATCTCTGGGTCATTTCTTTTATT 2397
Qy 550 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 569
Db 2398 ACATTTTCTTCTCAGGCTTCACATTTGGCAAGCTGGAGAGATCTCTCAACAGCGACTC 2457
Qy 570 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaIlePheAspGluLysGlu 589
Db 2458 CGATACATGGTTTCAATCCATGTCAGACAGGATGTGAGCTGGTTGATAACCTTAAA 2517
Qy 590 AsnSerThrGlyGlyLeuThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 609
Db 2518 AACACCATGGAGCATTGACCACCAAGCTCGCCACAGATGCTGCTCAAGTTAAGGGCT 2577
Qy 610 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629
Db 2578 ACAGAGCCAGACTGCTGTCTGTCATTACCCAGATATAGCAAAATCTTGGGACAGGAATCATC 2637
Qy 630 IleSerPheIleTyrGlyTyrGluMetPheLeuIleLeuSerIleAlaProValLeu 649
Db 2638 ATATCCCTAATCTATGGCTGGCAGTTGACACTTCTACTCTTAGCAATTTGCCCATCAT 2697
Qy 650 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 669
Db 2698 GCATAGCAGAGTGGTTGAGATGAAATGTTCTCTGGACAGCACTTAAAGATAAGAAG 2757
Qy 670 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 689
Db 2758 GAGCTAAGAGTTCTGGAAAGATTTGCTTCTGAAAGCAATAGAGAATTTCCGCACTGTGCTC 2817
Qy 690 SerLeuThrArgGluLysAlaPheGluGlnMetTyrCyluGluMetLeuGlnThrGlnHis 709
Db 2818 TCTTTGACTCGGAGCAGAAAGTTTGAATAATATGTATGCCAGAGCTTGCAGATACCATAC 2877
Qy 710 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 729
Db 2878 AGAATGCTCTGAAGAAAGCAGCGCTTTTGGGATCATCTTCTTCCATCCAGGCCATG 2937
Qy 730 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 749
Db 2938 ATGTATTTTTCGTATGCTGCTGTTCCGTTTGGTCTTACTTGTGTGGCGACGCGAATTA 2997
Qy 750 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 769
Db 2998 ATGACATTTGAAATGTTCTTAGTATTTCTCAGCTATTTGCTTTGGTGCATGGCAGTT 3057

QY 770 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 789
 DB 3058 GGCAGGCTGATTCATTGTCTGCTGATGATCCAAAGCAAGTGTGGCATCGCATATC 3117
 QY 790 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 809
 DB 3118 ATCATGATCATGAAAGAGTCCCTTCATTGACAGCTACAGCCGAGGAGCCCTGAAGCCT 3177
 QY 810 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 829
 DB 3178 AATACATTGGAAGGAATGTGAATTTAATGAAGTCTGTTCAACTACCCACCCGCGCT 3237
 QY 830 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 849
 DB 3238 GACATCCCATGCTGTCAGGGGCTGAACCTGGAGTGAAGAAGGCCAGAGCTGGGCCCTG 3297
 QY 850 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 869
 DB 3298 GTGGGAGCAGTGGCTGGGGAAGACACAGTGTCTCAGCTGTCCAGCGGTCTCAGAC 3357
 QY 870 ProValGlnGlyGlnValLeuPheAspGlyValAlaAlaLysGluLeuAsnValGlnTrp 889
 DB 3358 CCCATGCCGGAACAGTGTCTAGATGGCAAGAGTAACCACTAAATGTCCAGTGG 3417
 QY 890 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909
 DB 3418 CTCGAGCACACCTAGGATTTGTCTCCAGAGCCCATCTCTGTTGACTGCAGCATCGT 3477
 QY 910 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 929
 DB 3478 GAGAACATCGCTTATGGAGACACAGCCGGTCTGTCTCCAGATGAGATCGAGGGGG 3537
 QY 930 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 949
 DB 3538 GCGAAGGAAGCCAACTCCACAGTTCATTGACTGCTGCTGACAAATACACACAGA 3597
 QY 950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 969
 DB 3598 GTAGGAGACAAAGGACTCAGCTGTCCGGTGGGCGAGAAACAGCATCGCATAGCACGC 3657
 QY 970 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989
 DB 3658 GCCTCTGTGACAGAGCTCACAATTTGCTCTGATGAAGCAACATCAGCCCTGGATACA 3717
 QY 990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009
 DB 3718 GAGATGAAAGGTTGTCAGGAAGTCTTGGACAAAGCCAGAGAGCCGACCTGCATT 3777
 QY 1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1029
 DB 3778 GTGATCGCTCACCGCTGTCCACCATCCAGACCGCAGACTTGATCTGTGTGATTCAAGAT 3837
 QY 1030 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1049
 DB 3838 GCGAAGTCAAGAGCATGCGACCCACAGCAGCTGCTGCGACAGAAAGGCATCTATTTC 3897
 QY 1050 LysLeuValAsnAlaGlnSer 1056
 DB 3898 TCCATGCTGATGTGAGGCT 3918
 RESULT 8
 CRUPGP1185 3987 bp mRNA linear ROD 27-APR-1993
 LOCUS Chinese hamster p-glycoprotein mRNA (clone ADX185), complete cds.
 DEFINITION M59254
 ACCESSION M59254.1 GI:191156
 VERSION multidrug resistance glycoprotein; p-glycoprotein.
 KEYWORDS Chinese hamster, cDNA to mRNA, clone ADX185.
 SOURCE Cricetulus sp.
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae; Cricetulus.

REFERENCE 1 (bases 1 to 3987)
 AUTHORS Devine,S.E., Hussain,A., Davide,J.P. and Meiera,P.W.
 TITLE Full length and alternatively spliced pgp1 transcripts in multidrug-resistant Chinese hamster lung cells
 JOURNAL J. Biol. Chem. 266 (7), 4545-4555 (1991)
 MEDLINE 91154265
 PUBMED 1671863
 FEATURES
 source
 1..3987
 /organism="Cricetulus sp."
 /db_xref="taxon:10031"
 /cell_type="fibroblast"
 /tissue_type="lung"
 1..3987
 /gene="pgp-1"
 115..3624
 /gene="pgp-1"
 /codon_start=1
 /product="p-glycoprotein"
 /protein_id="AAA37005.1"
 /db_xref="GI:191157"
 /translation="MTTYYTYTGTGAGLVIVAYIQVSWFLAAGROIHKRQKPFHA
 IMNOEIGWFDVHDVGEINLRLTDDVSKINEGIDKIGMFQAMATFFGFIIGTRGW
 KLTIVLAISPVGLSAGIWAKILSSFTDKELQAYAKAGAAVEVLAIRTVIAFGSQ
 KKELERYNNLEBEAKRLGKKAITANISMGAAFLITIASYALAFWIGTSLVSKYSI
 GOVLTVPFAVLIAPISGQASPNTEAFANARGAAEYFNINDKPSIDFSKNGYKPD
 NIKGNLEPNHFSPSRKQVILKGLNLKVSQGTVALVGNKSGCKSTVQLQRLY
 CTGCVSISDGQDRTINVRYLREIIGVVSQEPVLFTTAENIRYGRNTVMDIEK
 AVKEGNAVDFIMKLPKHEPTLGERGAOLSGOKORAIARALVRNPKILLDEATSA
 LDTSEAVQAALDKAREGRTIIVAHRLSTVRNADIIAGDGGVIVFQGHHELMRE
 KGIYKLVMTQAGNEIELGNEVSGKNEIDNLMSKDSASSLIRRTSRRSIRGPH
 QDQRLKSTKEALDEDDPISFWRILKLNSEMPYFVVGIFCAIVNGALQPAFSTIFSK
 VVGVRNTDDETKRHSNLSFLILGVSIFTFQGGTFFGAGKILTKRLRYM
 FSKMLRQDVSWFNPKNTTGALTTRLANDAQVKATGARLAVITONTANTGTLIIIS
 LIYQWQLTLLALLAIVPIIAGVEMKLSQALKDKKELESGSKIAEIAENFTV
 SLTRQKFNMYASGLQIPYRNALKKAHVFGITFTQAMMIFSTAACFRFGALYAR
 ELMTFENVLLVFSIAIVGAMAVQSPADYAKAKVSASHIMIEKVPISDSYSTG
 GLKPTLEGNKVFNEVFNPTRPDIPLVGLNLEVKQGTALVSGSGCKSTVQL
 LERFYDPMAGTVFLDGKEVNQNLQWLAHLGIVSOEPLILFCIAENIAYGNSRVV
 SQDEIRAAKEANTHORIETSLPDKNTRVGGKGTQLSGOKORAIARALVRQPHILL
 LDEATSLDTESEKVVQELDKAREGRTCVIAHRLSTIQNADLIIVTQNGKVKHEGT
 HQQLLAQKGIYFSWVSQAGKR"
 BASE COUNT 1194 a 822 c 944 g 1027 t
 ORIGIN Map position 1q26.
 Alignment Scores:
 Pred. No.: 2,76e-196 Length: 3987
 Score: 3136.50 Matches: 598
 Percent Similarity: 75.82% Conservative: 211
 Best Local Similarity: 56.04% Mismatches: 215
 Query Match: 58.85% Indels: 43
 DB: 10 Gaps: 6
 US-09-873-409-4 (1-1058) x CRUPGP1185 (1-3987)
 QY 4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValaGluGluVal 23
 DB 496 TCATTACTGATAAAGAACTTCAGGCATATGCAAAAGCTGGAGCAGTCTCTGAAGAAGTC 555
 QY 24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43
 DB 556 TTAGCAGCCATCAGAACTGTGATTCATTGGAGGACCAAAAGAAAGAACTTGAA 609
 QY 44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63
 DB 609 -----
 QY 64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
 DB 610 -----AGGTACCAACCAATTTAGAAAGAGCCCAAGAACTTGGGATAAG 654
 QY 84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
 DB 84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103

Db 655 AAGCGATCAAGCCAAACATTTCCATGGGCGCAGCTTCTCTTATCTATGATCATAT 714
QY 104 GlyLeuAlaPheTptTyrGlyThrSerLeuLeuLeuAsnGlyGluProGlyTyrThrIle 123
Db 715 GCTCTGACCTTCTGGTATGGAACGTCCTTGGTCATCTCAAAAGAA-----TATCTATT 768
QY 124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
Db 769 GGACAAAGTGTCTCACTGTCTCTTGTGTATTAATTCGGCCATTCAGTATGGACAGCA 828
QY 144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163
Db 829 TCTCAAAATATCGAAGCCTTTGCCAATGCAAGAGAGCAGCCTATGAAATCTTCAACATA 888
QY 164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrIysProGluSerIle 183
Db 889 ATTGATTAATAGCCACGATTTGACAGCTTCTCAAGAAATGGGTACAAACCCAGACACATT 948
QY 184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
Db 949 AAGGAAATTTGGAATTCAAAATATATTCAGTTCAGTTACCCATCTCGAAAAGACGTTTCA 1008
QY 204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
Db 1009 ATCTTGAAGGGCCTCAACCTGAAGGTCGAGCGGACAGACAGTGGCCCTGGTTGGCAAC 1068
QY 224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243
Db 1069 AGTGGCTGTGGGAAAGACCACTGCTCCAGCTGCTGCAGAGGCTCTACGACCCACAGAG 1128
QY 244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
Db 1129 GCGTGTGCTAGTATCGACGGACAGACATCAGGACCATCAATGTAGGTATCTGCGGGAA 1188
QY 264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283
Db 1189 ATCAATGGGGTCGTAGTCAAGAACCTGTGTGTGTCCACCATATGCTGGAACATT 1248
QY 284 LysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303
Db 1249 CGCTATGGCCGAGAAATGTCACTATGCAGCAGATCAGAAAGCTGTCAAGGAAGCCAT 1308
QY 304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGly 323
Db 1309 GCCTATGACTTCATCATGAAACTGCCCCATAAATTTGACACTCTTGTGTGTGAGAGAGGA 1368
QY 324 AlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343
Db 1369 GCACAGTGTGTGAGGACAGAAACAGAGAAATCGCCATTCGTCGCCCTGGTCCGCAAC 1428
QY 344 ProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
Db 1429 CCCAAGATCCCTTTGTGGATGAGCGACATCAGCCTTGGACACAGAAAGTGAAGCCGTG 1488
QY 364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValAlaAlaHisArg 383
Db 1489 GTTCAGGCTGCTCTGGATAAGGCTAGAGAAAGCGGACCTACCATTTGTGTATGCCACCCG 1548
QY 384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403
Db 1549 TTGTCTACAGTTCGAAATGCTGACATCATAGCTAGCTGGGTGATGGTGTGGAG 1608
QY 404 LysGlyAlaHisAlaGluMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423
Db 1609 CAAGAAATCATGAAGAGCTCATGAGAGAGAGGAGGCAATTTACTTCAAACTTGTCATGACA 1668
QY 424 Gln-----AspIleLysLysAlaAspGluGlnMet 433
Db 1669 CAGACAGCAGGAAATGAAATTTAGGAAATGAAATGAGGTGTGTGAGTCTAAAAATGAAATT 1728
QY 434 GluSerMetThrTyrSerThrGluArgLysThrAsnSerLeu-----ProLeuHisSer 451
Db 1729 GATAAATTTAGACATGCTCTCAAAAGATTCAGCATCCAGTCTTAATCAGAGGAGATCAACT 1788

QY 452 ValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThr---GlnSerLys 470
Db 1789 CGCAGAAGTATCGTGGACCATGACATGACCAAGACAGGAAGCTTAGTACCAAGAGCGCTTG 1848
QY 471 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTyr 490
Db 1849 GATGAAGATGTATACCTCCAAATTTCTTTTGGCGTATCTTAAAGTTGAATTCATCTGAATGG 1908
QY 491 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 510
Db 1909 CCATATTTGTGGTGGTATATTTCTGTCATAGTAATAATGGAGCCTTGCACAGCATTC 1968
QY 511 SerIlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAspLysTyrThrLeu 529
Db 1969 TCAATAATATTTCTAAGGTGTGAGGGTTTTCACAAGAAATACTGATGATGAACCAAA 2028
QY 530 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 549
Db 2029 CGACATGATAGCAACTGTGTTTCTCTTATTTGTTTCTGATCTTGGGTCTATTTCTTTATT 2088
QY 550 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 569
Db 2089 ACATTTTCTTCAGGCTTCACATTTGGCAAGCTGGAGAGATCTCCACCAAGGCATC 2148
QY 570 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGluLysGlu 589
Db 2149 CGATACATGGTTTCAATCCATCTGTCGACACAGGATGTGAGCTGGTTTGATAACCCCTAAA 2208
QY 590 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 609
Db 2209 AACACCACTGGAGCATTTGACCACCAAGCTCGCAACGATGCTGTGTCAAGTTAAAGGGCT 2268
QY 610 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629
Db 2269 ACAGGACCAAGCTCTGTCATTTACCAAGATATAGCAAACTCTGGGACAGAAATCATC 2328
QY 630 IleSerPheIleTyrGlyTyrGluMetThrPheLeuLeuSerIleAlaProValLeu 649
Db 2329 ATATCCCTAATCTATGCTGGCAGTTGACACTTCTACTCTTAGCAATTTGCCCATCAT 2388
QY 650 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 669
Db 2389 GCATATAGCAGGATGTTGAGATGAAATGTTGCTCGCAAGCAAGCTATAAAGATAAGAAG 2448
QY 670 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 689
Db 2449 GAGCTAGAAGTTCTGGGAAGATTGCTCTGAAGCAATAGAGAACTTCGCACTGTGCTC 2508
QY 690 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 709
Db 2509 TCTTTGACTCGGAGCAGAGATTGAAATATGATATGCCAGAGCTTGCAGATACCATAC 2568
QY 710 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 729
Db 2569 AGAAATGCTCTGAAGAAAGCACACGCTTTTGGGATCACAATTCCTTCACCCAGGCCATG 2628
QY 730 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 749
Db 2629 ATGATATTTTCGATGCTGCTGTTTCCGGTTTGGTGTACTTGTGTGGCACCGGAATTA 2688
QY 750 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 769
Db 2689 ATGACATTTGAAATGTTCTATTAGTATTTCTCAGCTATTGCTCTTGTGGTCCATGGCAGTT 2748
QY 770 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaHisLeu 789
Db 2749 GGCAGGTGAGTTCATTTGCTCTGACTATGCAAGCAAGCAAGTGTGCGCATCGCATATC 2808
QY 790 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysPro 809
Db 2809 ATCATGATCATTTGAAAGATGCCCTTCCATTTGACAGCTACAGCAACCGGAGCGCTGAAGCCT 2868


```
QY 810 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 829
Db 2869 AATAACATTCGAAGAAATGTAATGAAGTCTGTGTCAACTACCCACCCGCGCT 2928
QY 830 AspValPheIleLeuArgGlyLeuSerIleGluArgGlyLysThrValAlaPhe 849
Db 2929 GACATCCAGTCTTCAGGGCTGAACCTGGAGGTGAAGAGGCCAGACGCTGGCCCTG 2988
QY 850 ValGlySerSerGlyCysGlyTyrSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 869
Db 2989 GTGGCAGCAGTGGCTGCGGGAAGACACACAGTGTCCAGCTGTCCAGCGGTTCTACGAC 3048
QY 870 ProValGlnGlyGlnValLeuPheAspGlyValAlaAlaLysGluLeuValGlnTyr 889
Db 3049 CCCATCCGCGGAACAGTGTTCAGTGGCAAGAGAGTAACCACTAAATGTGCCAGTGG 3108
QY 890 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909
Db 3109 CTCGAGACACCTAGGCATGTGTCCAGAGGCCATCTCTGTTCAGTCACATCGCT 3168
QY 910 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 929
Db 3169 GAGAACATCGCTATGGAGACACACCGCGGTCTGTCCAGCATGAGATCGAGAGGCG 3228
QY 930 AlaAsnAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 949
Db 3229 GCGAAGGAGCCCAACATCCACCATGTTCAATGAGTCACTGCTGCAACAATACACACAGA 3288
QY 950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 969
Db 3289 GTAGGAGACAAAGGACTCAGCTGTGGGTGGGAGAAACAGCATCGCCATAGCACGC 3348
QY 970 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989
Db 3349 GCCTCGTCAGACAGCTCACTATTTGCTCTGTGTAAGCAACATCAGCCCTGGATACA 3408
QY 990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009
Db 3409 GAGAGTGAAAGGTGTCCAGGAAGCTCTGGCAAAAGCCAGAGAGCCGACCTGTCAT 3468
QY 1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1029
Db 3469 GTATCGCTCACCCTGTCACATCCAGACGACACTGTATCGGTGGTATTCAGAAAT 3528
QY 1030 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuArgAsnArgAspIleTyrPhe 1049
Db 3529 GCGAAGTCAAGAGCATGCGACCCACAGCAGCTGCTGCGACAGAAGCATCTATTC 3588
QY 1050 LysLeuValAsnAlaGlnSer 1056
Db 3589 TCCATGCTCAGTGTGCAAGCT 3609

RESULT 9
CRUPGP1165 4304 bp mRNA linear ROD 27-APR-1993
LOCUS Chinese hamster p-glycoprotein mRNA (clone ADX165), complete cds.
DEFINITION M59253
ACCESSION M59253
VERSION M59253.1 GI:191154
KEYWORDS multidrug resistance glycoprotein; p-glycoprotein.
SOURCE Chinese hamster, cDNA to mRNA, clone ADX165.
ORGANISM Cricetulus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
Cricetulus.
REFERENCE 1 (bases 1 to 4304)
AUTHORS Devine,S.E., Hussain,A., Davide,J.P. and Melera,P.W.
TITLE Full length and alternatively spliced pgpi transcripts in
multidrug-resistant Chinese hamster lung cells
JOURNAL J. Biol. Chem. 266 (7), 4545-4555 (1991)
MEDLINE 91154265
PUBMED 1671863
FEATURES
Location/Qualifiers
source 1..4304

/organism="Cricetulus sp."
/db_xref="taxon:10031"
/cell_type="fibroblast"
/tissue_type="lung"
1..4304
/gene="pgp-1"
/cds="pgp-1"
111..3941
/gene="pgp-1"
/codon_start=1
/product="p-glycoprotein"
/protein_id="AAA37004.1"
/db_xref="GI:191155"
/translation="MEFEEDSGRKDKNLFKMGKSKKKEKKKDPVSVFTMFRVAGM
LDRLVLTGTLAAIHLGVALPLMLVDFGMDTDSFASVGNIPTNATNNATVQNASDIFG
KLEEMTYAYTIGAGVLIYAIQVSWFCAAGROHKIRKQIFGFMKLTFLVLA
DHDVDELNRLTDDVSKNEGIDKIGKFFQAMATFFGFTIGTFGKMLTFLVLA
SPVLGSAGIWAJLSSTFKLOQAKAGAAVEALAAIRTVIAFGQKLELYNN
NLEAKRIGIKKAITANTSMGAFLIYASVALAAMYGTSLVSKYSIGVLTVFFA
VLIAPPSIGQASPNIEAFANARGAAYEIFNIDNKPSIDSFSKNGYKPDNIKNLEPK
NIHFSYPSRKDVQILGKLNQVSGQYVALVGNSCGKSTTVQLLRLYDPTGCVVSI
DQDITRNRYLREIIGVSOEPVLFAITTAENIRYGRNVMTDEIKAVKBAAYD
FIMKLPHKFDLIVGERGAQLSGQKQRIAIARALVRNPKILLDLATSEAVV
QAALDKAREGRTTIVIAHRLSTVRNADIAGDGGVIVEQGNHSELMREKGIYFLVM
TOTAGNETELGNVESKNEIDNLDMSKDSASSLIIRRSSTRSIRGPHDQDKLSTK
EALDEDPVPSIFWRILKLNSEWPYVVGICAIYNGALQAPAFSIFSKVVGFTTNT
DDETRHDSNLSFLFLGLVSIFFELQGTGKAGEILLTKRUYRWFKSMURQDVL
SWFNPKNTGALTTRLANDAGVKGTARLAVITONIANLGTIIISLIYGWOLTL
LILATPTIATAGVVMKLSQALKDELEGSGKIATEATENFTVVSRLTREQKE
NMYAQLOIPIYRNALKKAHVGTITPFTQAMMYFSYACFRGALVARELMTFENVL
LVFSIAIVFGAMVGOVSSFPADYAKAKVSASHIMIIEKPSIDSYSTVGLKPNLTLEG
NVKNEVFNPTFRDIPVLOGLNLEVKKGQTLALVSGSGCKSTVQLLERFYDPM
GTVFKEVQNLNVQMLRAHLGIVSQPGLTSAENIATGDMNSRVVSQDEIBRAA
KEANTHQFEALDKYNTVRGDKGTQLSGQKQRIAIARALVRPHILLDLATSEALD
TESEKVKQIEMDKAREGRTTIVIAHRLSTIQNADLIIVVITQNGKVEHGHQQLAQKG
IVFMSVQAGAKR"
BASE COUNT 1275 a 895 c 1031 g 1103 t
ORIGIN Map position 1q26.

Alignment Scores:
Pred. No.: 3,04e-196 Length: 4304
Score: 3136.50 Matches: 598
Percent Similarity: 75.82% Conservativity: 211
Best Local Similarity: 56.04% Mismatches: 215
Query Match: 58.85% Indels: 43
DB: 10 Gaps: 6

US-09-873-409-4 (1-1058) x CRUPGP1165 (1-4304)
QY 4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluVal 23
Db 813 TCATTTACTGATAAAGAACTTCAGGCATATGCAAGCTGAGCAGTTGCTGAAGAAGTC 872
QY 24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuArgSer 43
Db 873 TTAGCAGCCATCAGAACTGTGATTCATTTGGAGGACAAAGAAAGAACTTGAA----- 926
QY 44 PheLeuLeuAsnIleThrArgTyrAlaTyrPheTyrPheProGlnTyrLeuLeuSerCys 63
Db 926 ----- 926
QY 64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
Db 927 -----AGGTACAAACAACAAATTTAGAGAAGCCAAAGAACTTGGGATAAAG 971
QY 84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
Db 972 AAAGCATCACCCCAACATTTCCATGGCGGCGAGCTTCTCTTATCTATGTCATCATAT 1031
QY 104 GlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrIle 123
Db 1032 GCTCTGGCCTTCCTGGTATGGAACGCTCTGGTTCATCTCAAGAA-----TATCTATT 1085
QY 124 GlyThrValLeuAlaValPhePheSerValIleHisSerTyrCysIleGlyAlaAla 143
```


[illegible]

QY 101 GlyThrTyrGlyLeuAlaPheTptTyrGlyThrSerLeuIleLeuAsnGlyGluProGly 120
DB 1054 GCATCATATGCCACTGGCCTTCTGGTATGGAATACCTTGATCTTAGCCCAATGAG----- 1107
QY 121 TyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIle 140
DB 1108 TACAGTATTGGCAATGTTCTTACAGTCTTTTCTGTACTGATTGGAGCTTTTCAGTATT 1167
QY 141 GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 160
DB 1168 GGACAGACTGCTCCAAGCATAGAGCATTTGCTTAATGCAAGAGAGCAGCCTATGCAATC 1227
QY 161 PheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysPro 180
DB 1228 TTTAATATCATACAAATGAACCTGAAATTTGACAGTTATTTCAGATCGAGGCCCAACCA 1287
QY 181 GluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgPro 200
DB 1288 GACCACATTAAAGGAATTTGGAATTCCTCAAAATGTTTCTTCAATTAACCCATCCAGACCA 1347
QY 201 SerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu 220
DB 1348 GATGTTGAGATCTGAAAGCCTTGAATCTCAAGGTTAAATTCGCGCAGACAGTAGCCCTG 1407
QY 221 ValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAsp 240
DB 1408 GTTGGTGGCAGCGCTGTGGGAAAGTACAACTGTTTCAGCTCATCCAGAGTTTACGAC 1467
QY 241 ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis 260
DB 1468 CCCAAGGAAGGCAGCATTAACCATTCAGCGGCAGGATCTGAAGAGCCTGAATGTAAGATAT 1527
QY 261 TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSer 280
DB 1528 CTGCGAGAGATTATTGTTGTGGTGAACAGGAGCCGCTGCTGTTGCTACAACTATTGCA 1587
QY 281 AsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArg 300
DB 1588 GAAATATTCTTATGGCCGTGAGGATGTCACCATGGAGAAATTTGAAAGCTGAGTACCA 1647
QY 301 GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly 320
DB 1648 GAAGCTAATGCTTATGATTTTCATCATGAAGCTCCCAAGAAGTTTGAATCTGGTTGGA 1707
QY 321 GluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu 340
DB 1708 GAAAGGCGGCACAGATGAGTGGAGGCGCAGAGCAGCGAATAGCAATGTCGCTGCTG 1767
QY 341 ValArgAsnProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspSerGluSer 360
DB 1768 GTTCACAAATCTTAAATTTCTGCTTGTATGAGGCAACGTCAGCTTTGGATACGTAGAGT 1827
QY 361 LysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValVal 380
DB 1828 GAATCCCTGTGCGAGCAGCAGCTGGCAAGCAGAGAGAGAGCAGCCACCGTTGTAGTA 1887
QY 381 AlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMet 400
DB 1888 GCTCATGCTCTATCAACAGTCCGAAATGCGATCTTATCGCTGTGTTGAAAGTGGAGTT 1947
QY 401 LeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeu 420
DB 1948 ATCAGAAACAAAGGAATTCATCTCAACTGATTTGAAAGGGAATCTCTATATAAATCT 2007
QY 421 ValMetSerGlnAspIleLysAlaAspGluGlnMetGluSerMetThrTyrSerThr 440
DB 2008 GTTAAATGAGACATAGAAATGAAATCCATCATCAGAAAAAGTGAATGCGATC 2067
QY 441 GluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPhe--- 459
DB 2068 AGTGTCAAAAGAGTGGATCTCAATCAATCTGGATGAATCCCTTAAATAAAGAAATTGAGA 2127
QY 460 -----IleAspLysAlaGluGluSer 466

DB 2128 AGAGGATCAACTAGGAGAGCATGAAAAAGCCAGAGAACCTTAATGATACCTGATGAGAA 2187
QY 467 ThrGlnSerLysGluLeuSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 486
DB 2188 GGTAGTTACCTGATGAAGAAATACCTCCAGTTTCATTTCTGAAACTATGAAAGTTGAAC 2247
QY 487 LysProGluTyrProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 506
DB 2248 AAAACGAATGCGCATCTTTGTAGCTGGAACCTTTCTGTGCAATTTGTCAATGGAGCTTGA 2307
QY 507 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLys 526
DB 2308 CAACCTGCATTTTTCAGTCATATTCTCAGAAATATATAGGAGATTTTTCAGAAAACTGATCAG 2367
QY 527 ThrThrLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIle 546
DB 2368 AAAGTTTGGAGAAAAAGAGCAACTTATATTCCTGCTGTTTTTTAGCACCTTGGATCAT 2427
QY 547 CysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThr 566
DB 2428 TCCTTTTTCATTTCTTGTTCAGGGTTTTGCAATTTGGCAAAAGCTGGAGAAATTTCTTACA 2487
QY 567 MetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaIleArgPheAsp 586
DB 2488 ATGAAGCTTCGATTTCATGCAATTTAAAGCAATGCTGAGACAGGACATGGCCTGGTTGAT 2547
QY 587 GluLysGluAsnSerThrGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIle 606
DB 2548 GATCTTAAATAATAGCACTGGAGCCTGACCAAGACTTGTCTAATGATCCCTCACAAAGTA 2607
QY 607 GlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeu 626
DB 2608 AAAGAGCAACTGGTCTCAGATTGGCAATTAATGCCCAAAACATAGCTAACCTGGGAAC 2667
QY 627 SerValIleIleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAla 646
DB 2668 GGAATATTATATCTTTAGTTTATGCTGGCAGCTGACCTGTTACTCTTAGCTGTGTGA 2727
QY 647 ProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLys 666
DB 2728 CCAATCATGCGATGCGCAAGAAATGATTGAAATGAAGATGTTGGCTGGCATGCCAAAAA 2787
QY 667 AspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArg 686
DB 2788 GATAAATAGAGCTGGAAGCTGCAGGAAAGATTGCCACAGAGCCATAGAAAAATATTCGA 2847
QY 687 ThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGln 706
DB 2848 ACTGTTGCTAGTTGACTCGGAAAAAAGATTGAGCTCATGTACGGAGAACATTTGCTT 2907
QY 707 ThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSer 726
DB 2908 GTACCATACAGAAATTTCTGTGAAGAGCGGCACATATTTGGGTTTTGTTTTCCCTTTCA 2967
QY 727 HisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGln 746
DB 2968 CAAGCAATGATGTTTTTACCTATGCTGCTGTTTCGCTTTGGAGCCTATCTTGTGTT 3027
QY 747 AlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAla 766
DB 3028 AATGGACACATAGATACAAACCTGTTTGTAGTGTGTTTTCAGCTGTTGTTTGTGCTG 3087
QY 767 MetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAla 786
DB 3088 ATGCGAATAGGACAAACCCAGCTCTTTTGACCAAGATATGCCAAAGCCAAAGATATCAGCA 3147
QY 787 AlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGly 806
DB 3148 GCCACTTGTGTTGCTGTTTAAATAGAGTACCTCCATAGACATGATACCGGAGGATGGA 3207
QY 807 LysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrPro 826
DB 807 LysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrPro 826

Db	3208	GAGAAACCTGAAAAATTTGGGGGAAACACACAGGATCAAGGATGTAATTTAACTATCCA	3267
Qy	827	CysArgProaspValPheIleLeuArgGlyLeuSerIleuSerIleGluArgGlyLysThr	846
Db	3268	AACCGACGAGGTCAAAATCCTCCAAGGCTTGAATCTAGCAGTAGAAAAAGGAGAAACG	3327
Qy	847	ValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArg	866
Db	3328	TTGGCCCTTGTGGTAGCAGTGGATGGAAAGAGCAGCTGTTCTCAGCTACTTGAGAGA	3387
Qy	867	LeuTyAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsn	886
Db	3388	TTTTATGATCCACTCAGTGGAGAAATTTGTTGATGACATGATGCAAGACACATAAT	3447
Qy	887	ValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCys	906
Db	3448	ATCCAGTGGCTGAGATCTCATATCGGTATCGTCTCACAGAGCCAACTCTGTTTCGACTTC	3507
Qy	907	SerIleAlaGluAsnIleAlaTyArgAspAsnSerArgValValProLeuAspGluIle	926
Db	3508	ACCATCGCTGAAACATCGGCTACGGGACACAGTCGGGAGGTGCCCATGAGGAAATC	3567
Qy	927	LysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTy	946
Db	3568	ATTAGTGCAGCGAAGCAGCAGTATTCATCTTCACTGACTCTCTGCCAGAAAATAC	3627
Qy	947	AsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAla	966
Db	3628	AATACTCGTGTAGGACAGCAAGGACACAGCTTCTGTTGTTGTTGTTGTTGTTGTTG	3687
Qy	967	IleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuLeuAspGluAlaThrSerAla	986
Db	3688	ATTGCTCGAGCTTATACGTAAAGCCCAAGATCTGCTACTGATGAAGTACATCTGCC	3747
Qy	987	LeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArg	1006
Db	3748	TTGGACAGACAGAAAGATGTTCCAGAGCAGCTGATTAAGCCAGAGGTGCG	3807
Qy	1007	ThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValVal	1026
Db	3808	ACCTGATCGTGATGCTCACCGCTGCTCCACCATCCAGATGCTGCAAGATGCTGTG	3867
Qy	1027	LeuHisAsnGlyLysIleLysGlnGlnGlyThrHisGlnGluLeuLeuArgAsnArgAsp	1046
Db	3868	ATCCAGAACGGCAAGCTCATAGACCAAGGAGCTCACCCAGCAGCTCTGCTGGAAGG	3927
Qy	1047	IleTyPheLysLeuValAsnAlaGlnSer	1056
Db	3928	TTCTACTCTGCTGGTCAATGTTCAAAGT	3957
RESULT 11			
LOCUS	OUA78609	3858 bp	linear
DEFINITION	Ovis aries multidrug resistance protein-1 (mdrl) mRNA, complete cds.		
ACCESSION	U78609		
VERSION	U78609.1	GI:2149086	
KEYWORDS			
SOURCE	Ovis aries.		
ORGANISM	Ovis aries		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Caprinae; Ovis.		
REFERENCE	1 (bases 1 to 3858)		
AUTHORS	Longley, M., Phua, S.H., van Stijn, T.C. and Crawford, A.M.		
TITLE	Isolation and mapping of the first ruminant multidrug resistance gene		
JOURNAL	Anim. Genet. 30 (3), 207-210 (1999)		
MEDLINE	99371931		
PUBMED	1042984		
REFERENCE	2 (bases 1 to 3858)		
AUTHORS	Longley, M. and Crawford, A.M.		
TITLE	Direct Submission		

JOURNAL	Submitted (18-NOV-1996) Biochemistry Department, University of Otago, Dunedin, New Zealand		
FEATURES	Location/Qualifiers		
source	1..3858		
	/organism="Ovis aries"		
gene	/db_xref="taxon:9940"		
CDS	1..3858		
	/gene="mdrl"		
	1..3858		
	/gene="mdrl"		
	/codon_start=1		
	/product="multidrug resistance protein-1"		
	/protein_id="AAB58489.1"		
	/db_xref="GI:2149087"		
	/translation="MDLEGDRNGRAAGNFKRDKRRFFSKDEKKEKRPVSTTFMF		
	IRSNWDLRLCMLVGLTAAI IHGAGLP LAMLVFGMDTDFAGAGNLGNITLSINSTST		
	RYDTEYGLKLEKEMTYAYYYGIGAGVLIAYIOVSFWCLAAAGQVHRI RKQFFHAI		
	MQEELGWFDVHDVGLNTRLNDYSKINEGIDKIMFQPMATFTFCFTIGFTGWN		
	LTVLIALIPVLGSLAAIWAKLSFTDKLLAYAKAGAAVEVLAATVIAFGQK		
	KELERYNKLEBEAKRIGIKKAITANISMAAPFLLIYASALAFWYGTSLVLSREYSIG		
	QVLTVPFSLVIGAFSIGQASPNIEAFANARGAEVFKIIDNKPIDSYSTNGHKPDN		
	IKGNLEFRNVHFHPSRNEVKILKGLNLKVGSGQTVLVNCGGCKSTTVQLMQRLYD		
	PTGEMVSDIGQDIRTINVYLRLEIIGVVSQEPVLPAITAEINIRYGRDVTWDELOKA		
	VKEANAYDFIMKLPNKFTLVGERAQLSGGKORIIATARALVRNPKILLDEATSA		
	GTESEAVQAALDKARKGRTIIVIAHRLSTVRNADIIAGLDGCVIVEGSHDELMKR		
	GLYFKLVTMTQNGNELELETPGESLKTDDLYMSSQDSRSRLIRKSTRRSIRGSQS		
	QDKRLSTEETLDESVPVSRFWRLKINITEWPFYVGVFCALINGALOPAFSVIFSR		
	IGIFPRNDNETKRONSLFLLLILGIIISFTFLOQTFGKAGELITRLRLVLF		
	RSMLRDYVSFDDPKNTTGALTTLANDAAQVKGAVGSLAVITQNIANLGTGIIISL		
	IYGWQLTLLALVPIIIVAGVIEMLSGQALKDKKELEGAGKATATEINERFVVS		
	LTREERFYTAQSLQVPRNSLRRAHVFGITFSIQAMFYFSYAGCFRFGAYLVAQG		
	IMEFDVILLVFSVAVGVAMQVSTFAPDYAKAKVSAAHVINIIEKPLIDSYSTEG		
	LKPSTVEGSAFNDVFNYPTRPDVPLRGSLVLEKGTQLALVSSGCGKTVVQLL		
	ERFYDLACTVFDGKEVKQLNVMRAHMGIVSOEPIILFDCSICENIAYGDNVVS		
	QEEIHAKEANIHSFIEMLPKYKTRVGDGTQLSGGKORIIATARALVRPHILL		
	DEASALDTESEKVVQEAELDKAREGRTCTVIAHRLSTIQNADLIIVFQNGRIKEGHT		
	QQLAAQGLIYFTMVSVQAGTKRQ"		
BASE COUNT	1072 a 853 c 1002 g 931 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	4,01e-195	Length:	3858
Score:	3118.50	Matches:	602
Percent Similarity:	75.30%	Conservative:	203
Best Local Similarity:	56.31%	Mismatches:	217
Query Match:	58.51%	Indels:	47
DB:	4	Gaps:	6
US-09-873-409-4 (1-1058) x OUA78609 (1-3858)			
Qy	4	SerLeuThrSerLysGluLeuSerAlaTySerLysAlaGlyAlaValaGluGluVal	23
Db	727	TCATTTACTGATAAAAACTTTTGGCATATGCAAAAGCTGGAGCAGTAGCTGAAGAGT	766
Qy	24	LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer	43
Db	787	TTAGCGGCTATTAAAACTGTGATTGTCATTGGAGGACAAAAAAGAACTTGAA	840
Qy	44	PheLeuLeuAsnIleThrArgTyAlaTrpPheTyPheProGlnTrpLeuLeuSerCys	63
Db	840	-----	840
Qy	64	ValLeu**PheValArgTyThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys	83
Db	841	-----AGGTACACAAAACTTGGAGAGCCAGAGAAATTGGAATAAG	885
Qy	84	ArgThrIleAlaSerLysValSerLeuGlyAlaValTyPhePheMetAsnGlyThrTy	103
Db	886	AAGGCTATCAGCGCCCAACATCTCTATGGAGCTGCTTTCTTGTGTATCTATGATCCT	945
Qy	104	GlyLeuAlaPheTrpTyGlyThrSerLeuIleLeuAsnGlyGluProGlyTyThrIle	123
Db	946	GCCTGGCGTTCTGGTACGGGACCTCTCTTGGTCTCTCAAGAGAA-----TATTCAT	999


```
QY 124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
DB 1000 GGCAAGTGTCTACAGTCTCTTTCTGTATTAAATGGAGCTTTAGTAGTGACAGGCA 1059
QY 144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaPheHisIlePheGlnVal 163
DB 1060 TCTCCAAACATTGGAGCATTTGCAACCGCAAGAGGAGCAGCTTATGAAGTCTTCAAGATC 1119
QY 164 IleAspLysIleValProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
DB 1120 ATTGATAACAAACCAAGCATCTGACAGCTATTCAAACACTGGGCACAAACCCGCAATATT 1179
QY 184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
DB 1180 AAGGAATCTGGAATTCAGAAATGTTCACTTCCATTAACCATCTCGAACCAAGATTAG 1239
QY 204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
DB 1240 ATCTGAAGGCGCTCAACCTGAAGTAGGAGCGGCGAGCGGTGGCCCTGGTTGGGAAC 1299
QY 224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243
DB 1300 AGTGGCTGTGGGAAGACACACCGTCCAGCTGATGACAGAGGCTGTATGACCCCGAG 1359
QY 244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
DB 1360 GGATGCTAGTATCGACGACAGGATATCAGGACCATCAACGTAAAGTATCTCGGGAG 1419
QY 264 HisIleGlyValValSerGlnProValLeuPheGlyThrThrIleSerAsnAsnIle 283
DB 1420 ATTATCGGGTGTGAGTCAGGAGCGCTGTCTGTGTGCCACCATGATGAGTGTGAGAAT 1479
QY 284 LysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303
DB 1480 CGTACGCGCGTGAAGATGTCACCATGGATGAGATTGAGAAAGCTGTGAAGGAAGCCAAC 1539
QY 304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323
DB 1540 GCCTATGACTTTATCATGAACCTGCCTTAATAGTTTGACACCCCTGGTTGAGAGAGGA 1599
QY 324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343
DB 1600 GCCCAGCTAAGCGGTGACAGAGCAAGAAATCGCCATTGCCCGGCGCTGTGTCGCAAC 1659
QY 344 ProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
DB 1660 CCCAAGATCTCTGTTGGATGAGGCCACCTCAGCGCTGGGCCACCGAGAGTGAAGCGGTG 1719
QY 364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyValArgThrThrIleValAlaHisArg 383
DB 1720 GTTCAGCGCGCCCTGGATAGGCGAGAAAGCGGACCCACCATTTGTATGATGCGCATCGT 1779
QY 384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403
DB 1780 CTGTCCACGGTTGCAATGCTGATCATCTGCTGCTTGTATGATGAGTCAATCTGTCAACATG 1839
QY 404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423
DB 1840 GAAGGCGAGTCATGATGAGCTCATGGGAAAGAGGCAATTTACTTCAAACTGTGCACAAATG 1899
QY 424 GlnAspIleLysLysAlaAspGluGluMetGluSerMetThrTyrSerThrGluArgLys 443
DB 1900 CAG---ACAAAAGGAATGAACTTGAATTTGGAATAATCCCTGGTGAATCCCTAAGTAAA 1956
QY 444 ThrAsnSerLeuProLeuHisSerValLysSerIleLysSerPheIle----- 460
DB 1957 ACTGATGCTGTACATGCTCCTCCCAAGATTCA---AGATCCAGTCTTAATCAGAAGAAAA 2013
QY 461 -----AspLysAlaGluGluSerThrGln 468
DB 2014 TCAACTCGCAGGAGTATCCGTGGATCCCAAGCCAGGACAGAAAGCTCAGTACAGAGAA 2073
```



```

Db 3154 CGACAGACGTCCTCGAGTGTCTGGGGGCTGAGCCTGGAGGTGAAGAAGGCGCCAGACGCTG 3213
Qy 848 AlaPheValGlySerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeu 867
Db 3214 GCGCTGTGGCAGTAGTGGCTGTGGAGAGACACGGTGTCCAGCTCTCGAGCGGTTC 3273
Qy 868 TyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGlnLeuVal 887
Db 3274 TATGACCCCTTGCTGGCAGACAGTGTATTGACGGCAAGAGTAAGCAGCAGTGAATGTC 3333
Qy 888 GlnTrpLeuArgSerGlnLeuAlaLysValProGlnGluProValLeuPheAsnCysSer 907
Db 3334 CAGTGGCTCGGGGCACATATGGCATCGTGTCCAGAGGCCATCTGTTGACTGCAGC 3393
Qy 908 IleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluLysLys 927
Db 3394 ATCGGGAGAAATATCCCTACGGAGACAACAGCCGGTGTGTCCAGAGAGATGAG 3453
Qy 928 GluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsn 947
Db 3454 CACGGGGCCCAAGAGGCCAACATCCACTCTTCATCGAGATGCTGCCGACAAATACAAC 3513
Qy 948 ThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaLys 967
Db 3514 ACCAGAGTAGGACAAAGGAATCAGCTCTCTGTTGGCCAGAAACAGCGCATGTCATA 3573
Qy 968 AlaArgAlaLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeu 987
Db 3574 GCTCGGCCCTTGTAGACAGCTCACTCTGCTCTGTGATGAAGTCACTCGGCTCTG 3633
Qy 988 AspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThr 1007
Db 3634 GACACGGAAGTGAAAGTGGTCCAAAGAGCCCTGGACAAAGCCGAGAGGACGAC 3693
Qy 1008 CysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuLeuValLeu 1027
Db 3694 TGCATCGTATCGCTCACCAGCTGTCCACCATCCAGACGACAGTGTAGTGGTGT 3753
Qy 1028 HisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAspIle 1047
Db 3754 CAGATGGCAGAAATCAAGAGCATGTCACACACAGCAGCTGTGCGGCAAGAGCATC 3813
Qy 1048 TyrPheLysLeuValAsnAlaGlnSer 1056
Db 3814 TATTCACCATGTCAGTGTCCAGGCT 3840
RESULT 12
AF257746
LOCUS Rattus norvegicus multidrug resistance protein 1a (Pgyl) mRNA,
DEFINITION complete cds.
ACCESSION AF257746
VERSION AF257746.1 GI:7739772
KEYWORDS Rattus norvegicus.
SOURCE Rattus norvegicus.
ORGANISM Rattus norvegicus.
REFERENCE 1 (bases 1 to 4927)
AUTHORS Hooiveld,G.J.E.J., Wilms,J.W.J., Hagenbuch,B., Jansen,P.L.M.,
Meijer,D.K.F. and Muller,M.
TITLE Cloning and functional characterization of the rat multidrug
resistance protein Mdr1a
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 4927)
AUTHORS Hooiveld,G.J.E.J., Wilms,J.W.J., Hagenbuch,B., Jansen,P.L.M.,
Meijer,D.K.F. and Muller,M.
TITLE Direct Submission
JOURNAL Submitted (19-Apr-2000) Div. Gastroenterology & Hepatology,
University Hospital Groningen, Hanzeplein 1, Groningen NL-9713 GZ,
The Netherlands

```

```

FEATURES
source
Location/Qualifiers
1..4927
/organism="Rattus norvegicus"
/strain="Wistar"
/db_xref="taxon:10116"
/sex="male"
/cell_type="hepatocyte"
/tissue_type="liver"
/dev_stage="adult"
1..4927
/gene="Pgyl"
86..3904
/gene="Pgyl"
/notice="ATP-binding cassette transport protein; Abcb1a"
/codon_start=1
/product="multidrug resistance protein 1a"
/protein_id="AAF69007.1"
/db_xref="GI:7739773"
/translation="MELEEDLNGRADKNFSGKMGKSKKEKKEKPAVSLTMRVAGM
LDRLYMLLGLTAAIHHGIALPLMLVFGDMTDSFANVGNRMSFYFNATDIYAKLEDE
MTTAYYYTGTGAGVLIVAYTVQSLWCLAGRQIHKQKFFHAIMNQEIWMFVDHV
GELNRLTDDVSKINIGIDKIGMPFQAMATFGGFIIGTRGKWLTVILVLAISVLG
LSAGIWKILSSFTDKELQAKAGAAVEVLAARTVIAFGQKKELENNLEEA
KELGKALTANI SMGAFLIIVASALAFVGTSLVLSKEYTIGQVLTVPFSLVIGA
FSVQASPNIEAFANRGAAYEVSIDNKPISIDSPSKCHKPONI QGNLFKNHFS
YPSRDVOILKGLNKVKSGTVALVNSGCGKSTTVQLLQRLDPIEGEVSIDGDI
RTNRYLREITIGVVSQEPVLFPATIAENIRYGRNVTMDIEKAKKANAYDFIMKL
PKHPTLVGERGAQLSGGQKORIAIARALVRNPKILLDEATSLDTESEAVVQALD
KARGRTTIVIAHRLSTVRNADVAPDGVIVVEQGNHDLMRKGIYFKLVMTQTAG
NEILGNEACESKDQIDNVMSKDSGSLIRRRSTRKIRGPHQDGLSTKEALDD
DVPASFWRLKLNSTWPFYVFCALINGLQPAFSIIFSVKVGVTKNDTPEIQ
RONSPLFLFLIIGTIFTFQGGTGTGAGETLTKRLRMVFKSMLRQDISWFD
PKNTGALTTLRLANDAAQKATGSRVAVITQNIANGTGIISLIQWQLTLLLA
VPIITAGLVEMKMLSGQALKKELEGKIATEAENFRVTLSLTREQKPTMYAQ
SLOIPYRNALKKAHVGTFTSOTAMMYFVSACFRFAYLVARELMTFENVLLVFA
IVFGMAVGVSFPADYAKAKVSASHIIRIIEKIPEIDSVSTEGCLKMNLGKFN
GMVNYPTRPNIPLVGLSLEVKGQTLALVSSCGSKSTVQLLREYFDPNAGTVEL
DQKEIKQLNVQWLRHLGIVSOEPLIFDCSIAENAYIGNRSVRVSHBEIVKAAKANI
HGKIDSLPEKYNTRVGDGTQSGQKORIAIARALVRNPKILLDEATSLDTESEBK
VVOBALDKAREGRTCTIVIAHRLSTIQNALDILVVIQNGQVKHGHTHQQLLAQKGIYFSM
VSVOAGAKRS"
BASE COUNT 1474 a 1037 c 1190 g 1226 t
ORIGIN
Alignment Scores:
Pred. No.: 7,34e-195 Length: 4927
Score: 3116.50 Matches: 601
Percent Similarity: 75.73% Conservative: 207
Best Local Similarity: 56.33% Mismatches: 216
Query Match: 58.47% Indels: 43
DB: 10 Gaps: 7
US-09-873-409-4 (1-1058) x AF257746 (1-4927)
Qy 4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValaGluGluVal 23
Db 773 TCATTTACTGATAAGAACTCCAGGCTTATGCAAAAGCTGGAGCAGTTCGCTGAAGAATC 832
Qy 24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43
Db 833 TTAGCAGCCATCAGAACTGTGATTCCTTTTGGAGGACAAAGAAAGAACTTGAA----- 886
Qy 44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63
Db 886 ----- 886
Qy 64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysaspPheGlyIleLys 83
Db 887 -----AGGTACAATAACAATTTGGAAGAAGCTAAAGAGGCTTGGGATAAAG 931
Qy 84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
Db 932 AAAGCTATCACGCGCAACATTTCCATGGGTGGCAGCTTTTTCGTTATCTATGTCATCAT 991

```


QY 104 GlyLeuAlaPheTrpTyrGlyThrSerLeuLeuLeuAsnGlyGluProGlyTyrThrIle 123
Db 992 GCTCTGCATCTGGTATGGACTTCCTTGGTCATCTCAAGAANA-----TACACTATT 1045
QY 124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
Db 1046 GGCAAGTGTCTCACTGTCTCTTTTCTGTATTAAATTGGAGCAGTCAGTGTGGCAGGCA 1105
QY 144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163
Db 1106 TCTCCAAATATTGAAGCTTCGCGCAATGCTAGAGGAGCAGCTTATGAAGTCTTCAGTATA 1165
QY 164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
Db 1166 ATTGATTAATAGCCCATATAGACAGCTTCTCAAGAGTGGGCACAAACCCGACACATA 1225
QY 184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
Db 1226 CAAGGAAATTTGGAAATCAAAAATATTCACTTCAGTTACCCGCTCTCGAAAAGACGTTTCAG 1285
QY 204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
Db 1286 ATCTGAAGGCGCTCAACCTGAAGGTGAAGCGGAGCGGAGCGGTAGCCCTGGTTGGCAAC 1345
QY 224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243
Db 1346 AGTGGCTGTGGAAAGACACAACTGTCTCAGTCAGTGTGGAGGGCTCTACGACCCCATAGAG 1405
QY 244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
Db 1406 GCGAGGTCAGTATCGACGCGACAGGACATCAGGACCATCAATGTAGGTATCTCGCGGAA 1465
QY 264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283
Db 1466 ATCAATCGGGTGTGAGTCAGGACCGCTGTGTTCGCCACCAATTCGCCGAAACATT 1525
QY 284 LysTyrGlyArgAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsn 303
Db 1526 CGTATGCGCGAGAAACGTCACCATCGATGATGAGTAGAGAAAGCTGTCAAGGAAGCCAT 1585
QY 304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyLysGly 323
Db 1586 GCCTATCATTTCAATGAAACTGCCCCACAAATTTACACACCCCTGGTGTGGTAGAGAGGG 1645
QY 324 AlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343
Db 1646 GCGCAGCTGAGTGGGGACAGAAACAGAGGATCGCCATTCGCCGGGCCCTGGTCCCGAAC 1705
QY 344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
Db 1706 CCCAAGATCCTTTTGGATGAGGCCACGTCAGCTTGGACACAGAAAGCGGAGCGGTG 1765
QY 364 ValGlnAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383
Db 1766 GTTCAGGCCGCTCTGGATAAGGCTAGAGAGCGCGGACCACTTGTGTATGATGCTACCCG 1825
QY 384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysaspGlyMetLeuAlaGlu 403
Db 1826 TTGTCTACAGTTCGCAATGTGACGTCATTCGTTGGTTTGTATGTTGATGTTGATGGAG 1885
QY 404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423
Db 1886 CAAGGAATCATGATGAGTCTCATGAGAGAAAGAAATTTACTTCAAACTGTGTCATGACT 1945
QY 424 Gln-----AspIleLysLysAlaAspGluGlnMetGluSer----- 435
Db 1946 CAGACAGCAGGAAATGAAATGAAATAGGAAATAGGAAATCAAGCTTGTGAATCTAAAGATGGAATT 2005
QY 436 -----MetThrTyrSerThrGluArgLysThrAsnSerLeu-----ProLeuHisSer 451
Db 2006 GATAATGTGGACATGCTCTCAAAAGATTTCAGGATCCAGTCTTAATGAAGAAAGATCAACT 2065

QY 452 ValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThr---GlnSerLys 470
Db 2066 CGCAAAAGCATCGTGGCCACCATGATCAAGACGGGAACTTAGCACCAAGAGGCTCTG 2125
QY 471 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTyr 490
Db 2126 GATGACGACGTACTCTCCAGCTTCTTTTGGCGGATCCTGAAGTTGAATCAACTGAATGG 2185
QY 491 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 510
Db 2186 CCTATTTTGTGGTTGGTGTATTTTGTGTCATTAATAGGAGGCTTGCACAGCATTC 2245
QY 511 SerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThrLeu--- 529
Db 2246 TCCATAATATTTCAAAAGGTGTGAGGGTTTTTACAAAAAATGACACCCCTGAAATCCAG 2305
QY 530 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 549
Db 2306 CGCAGAACAGCAACTTGTTTTCTTTATTTGTTCTGATCCTTGGGATCATCTCTTTCAAT 2365
QY 550 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 569
Db 2366 ACCTTTTCTCTCAGGCTTCACATTTGGCAAGCTGGAGAGATCTCTCACAAGCAGCTC 2425
QY 570 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaIleArgPheAspGluLysGlu 589
Db 2426 CGATACATGCTTCAATCCATGCTGAGACAGGACATAGCTGGTTGATGACCCCTAAA 2485
QY 590 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 609
Db 2486 AACACACAGGAGCGCTGACCACAGGCTTGCCAATGACGCTCTCAAGTGAAGGGCT 2545
QY 610 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629
Db 2546 ACAGGTCTAGGCTTCTGTTATTACCCAGAACATAGCAAAATCTTGGACAGGCATC 2605
QY 630 IleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeu 649
Db 2606 ATATCCCTGATCTACGCTGGCAATTGACACTTTTACTCTAGCAATTTGTTCCCATCAT 2665
QY 650 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 669
Db 2666 GCATATAGCAGGAGTGTGAAATGAAATGTTGTCTGGACAGCGCTGAAAGATAAGAG 2725
QY 670 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 689
Db 2726 GAACAGAGTCTCGGAGAGATCGCTACAGAGCAATTTGAAAACTTCCGACTGTCGCT 2785
QY 690 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 709
Db 2786 TCTTTGACTCGGAGCAGAGAAATTTGAAACTATGATGCCAGAGCTTGCAGATACCATAC 2845
QY 710 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 729
Db 2846 AGAAATGCTTTGAAGAAAGCGCACGCTTGTGGATCAGCTTCTCTTCCCTCACCCAGGCCATG 2905
QY 730 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 749
Db 2906 ATGTATTTCTCTATGCTGCTTGTTCGGTTTGGTTCCTTCTTGTGGCGCAGAACTC 2965
QY 750 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 769
Db 2966 ATGACATTTGAAATGTTCTGTTAGTATTCTCAGCTATTGCTTTGTTGGTCATGGCAGTG 3025
QY 770 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaHisLeu 789
Db 3026 GGCAGGCTGAGTTTCAATTCGCTCTGACTACGCGAAAGCCAAAGTCTCGGCATCCCAATC 3085
QY 790 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 809
Db 3086 ATCAGGATCATTTGAGAAATCCCTGAGATTGACAGCTACAGCAGGAGGCTTGAAGCCT 3145
QY 810 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPheTyrProCysArgPro 829

Db	3146	AAATATGTTGAAGGAATGTGAATTTAATGGAGTCATGTTCAACTATCCACCCGACCC	3205
Qy	830	AspValPheileuArgGlyLeuSerLeuSerileGluArgGlyLysThrValAlaPhe	849
Db	3206	AACATCCAGTGCTTCAGGGCTGAGCTAGAGGTGAAGAAAGGCAGCGCTGCCCTC	3265
Qy	850	ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp	869
Db	3266	GTGGGACAGTAGTGGCTCGCGGAAGAGTACAGTGGTCCAGCTGCTTGAGCGCTTCATGAC	3325
Qy	870	ProValcInGlycInValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp	889
Db	3326	CCCATGGCCGAACAGTGTTCAGATGGCAAGAAATAAAGCAACTCAATGTCAGTGG	3385
Qy	890	LeuArgSerGlnileAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla	909
Db	3386	CTCGGGCCACCTGGGCATTTGTGCCAGGAGCCATCTCTGTTGAGTCGAGCATGCC	3445
Qy	910	GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla	929
Db	3446	GAGAAATTCCTTACCGAGACACAGCCGTGCTGTCTCATGAGGAGATCGTGAAGGCA	3505
Qy	930	AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln	949
Db	3506	GCCAAAGGAGGCCAACATCCACCATGTTTCATCGACTCACTGCCCTGAGAAATACACACAGA	3565
Qy	950	ValGlyLeuLysGlyValaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg	969
Db	3566	GTGGGAGACAAAGGACTCAGCTGTGCGCGGGCAGACAGCGCATGCCATCGCGGC	3625
Qy	970	AlaLeuLeuGlnLysProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspAsn	989
Db	3626	GCCTCGTCAGACAGCCTCACATCTTACTTCTGGATGAAGCGACATCAGCTCTGATACG	3685
Qy	990	AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu	1009
Db	3686	GAGAGTGAAGAGTCTGTCAGGAAGCGCTGGGACAAAGCCAGGAGGCGCACCTGCATT	3745
Qy	1010	ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn	1029
Db	3746	GTGATCGCGCACCGCTGTCCACATCCAGAACGAGACTTGATCGTGTGATTCAGAAC	3805
Qy	1030	GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe	1049
Db	3806	GGCAGGTCAAGGAGCAGCGACCCACACAGCAGCTGCTGGCCAGAAAGGCATCTATTTC	3865
Qy	1050	LysLeuValAsnAlaGlnSer	1056
Db	3866	TCGATGTCAGTGTGCAGGCT	3886
RESULT	13		
AX105082			
LOCUS	AX105082	4279 bp	DNA
DEFINITION	Sequence 26 from Patent WO0123540.		linear
ACCESSION	AX105082		
VERSION	AX105082.1		
KEYWORDS	GI:13921232		
SOURCE			
ORGANISM	dog.		
REFERENCE	Canis familiaris		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomom		
TITLE	Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.		
JOURNAL	1 (bases 1 to 4279)		
FEATURES	Stockner,P.J., Steimel-Crespi,D.T., Crespi,C.L., Reif,T.C. and		
source	Patten,C.J.		
	P-glycoproteins and uses thereof		
	Patent: WO 0123540-A 26 05-APR-2001;		
	GENTEST CORPORATION (US)		
	Location/Qualifiers		
	1. .4279		
	/organism="Canis familiaris"		
	/db_xref="taxon:9615"		
CDS	17. .3862		

```

/note="unnamed protein product"
/codon_start=1
/protein_id="CAC37723.1"
/db_xref="GI:13921233"
/translations="MDEGGKGRGSAEKNFWKMGKKKKKKKKKKKPTVSTFAMFRYSN
WLDRLVYLVTMAAIIHGALPLMLVFGNMTDSFANAGISRNKTPPVII NESITNNT
QHEINHLDEEBMTYATYYSGIGAGVLVAAYIQVSWFCLAAGRQLIKRKQFFHAIKMLQ
EIGWFDVHDVGEINLTDDVSKINEGIDGKIMGFQSIATFTFTGTVGTFGRGKWLKL
VILAIISPLVGLSAAIIWAKLISFTDKELLYAKAGAVAEVLAARIVIAFGGQKLT
ERYKNLEAGKIGIKKATITANIISGAFFLIYASVALAFWYGTSLVLSSEYTGQVL
TFPSVLNIGAFSISGOASPSITEAFANARGAAVEIFKIIDNKPSIDSKSKGHKPDNIG
KSGNKHVHSPSRKEVKLLKGLNWKVQSGOTVALVGNSSCGKSTTVOLMQRVLDPTD
GMVICDQDITRTINVRHLREITGVGQEPVLPATTTAENRYRGRENTWMEIEKAVKE
SEAYVQVADLKARQKGRITIVIAHRLSTVRNADVIAGPDDGVIKVEGNHDELMKEGVI
FLKVTWGTGNETELENATESKESDALEMSPKDSSSLIKRSTSTRSITHAPQGDQR
KLTGKBQDLNENVPVFWRLTLKUNSTPFPVVGIFCAIINGLOPAFSAIFSRIGI
FTDPEPTKRONSNFSLVFLGIIISFTFFLQGTFCGKAGELLKRLRYMVWRSM
LRQVSWFDDPKNTGALTRLANDAAQVKALGSLRAVTVQNIANLGTGIIISLIVG
WQUTLLLAITVPIIAIAGVEMKLSQALKDKKELEGAKIATEIENRTVSVLSR
EORPEYVIAFGVGRYNSRLKAHIFGVFSITQAMMYFSYAGCFRFGALVANEFN
FODVLVFSALVFAGMAAGVQSSFPADYAKAKVSAAHVIMIEKSPILDSYSHGLPR
NTLEGNTFNEVFNPTRPDIIVLQGLSLEVKKGTTLVSGSSCGKSTTVOLLRPE
YDLAGSVLIDGKEIKHLNVQWLRHLHGIUSQSPILFDCSIAENIAYGDNRSRVSHHE
IVQAKEANIIHFTPELXINTRVGDKTQSGQQRQIARARLVQRQPHILLDEEA
TSALDTESEKRVQEQALDKRGEGRCTIVIAHRLSTIONADLIVVFQNGKVKHEGTHQOL
LAQGIYFSMVSVQAGAKR"
BASE COUNT      1296 a      833 c      1009 g      1141 t
ORIGIN
Alignment Scores:
Pred. No.:      7,17e-195      Length:      4279
Score:          3115.50      Matches:      602
Percent Similarity: 75.02%      Conservative: 200
Best Local Similarity: 56.31%      Mismatches: 220
Query Match:      58.45%      Indels:      47
DB:              6      Gaps:      6

US-09-873-409-4 (1-1058) x AX105082 (1-4279)

Qy  4  SerLeuThrSerLysGluLeuSerAlaTyvSerLysAlaGlyAlaValaGluGluVal 23
Db  734 TCATTACTGATAAAGAACTCTTGGCCCTATGC AAAAGCTGGAGCAGTAGCTCGAAGAAGTC 793
Qy  24  LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43
Db  794 TTAGCAGCAATCAGACTGTGATTGCCCTTTGGAGGACAAAGAAAGAACTTGAA----- 847
Qy  44  PheLeuLeuAsnIleThrArgTyvAlaTrpPheTyvPheProGlnTrpLeuLeuSerCys 63
Db  847 ----- 847
Qy  64  ValLeu**PheValArgTyvThrGlnAsnLeuLysAspAlaLysAapPheGlyIleLys 83
Db  848 -----AGGTACAAACAAAATAATTGAAGAAAGCTAAAGGAATTTGGGATAAAG 892
Qy  84  ArgThrIleAlaSerLysValSerLeuGlyAlaValTyvPhePheMetAsnGlyThrTy 103
Db  893 AAAGTATACAGCGCAACATTTCTATTGGTGCCGCTTCTTATTGATCTATGCATCATAT 952
Qy  104 GlyLeuAlaPheTrpTyvGlyThrSerLeuIleLeuAsnGlyGluProGlyTyvThrIle 123
Db  953 GCTCTGGCTTTCTGTATGGGACCTCTCTTGGTCTCTCCAGTGAA-----TATACTATT 1006
Qy  124 GlyThrValLeuAlaValPhePheSerValIleHisSerTyvCysIleGlyAlaAla 143
Db  1007 GGACAAGTACTCACTGTCTCTTTTCTGTATTAATTTGGGCTTTTAGTATTGGACAGGCA 1066
Qy  144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163
Db  1067 TCCCCAAGCATTTGAAGCATTTGCAAAACGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1126
Qy  164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyvTrpLysProGluSerIle 183

```


Qy 1028 HisAsnGlyIysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIle 1047
Db 3761 CAGATGGCAAGTCAAGGAGCATGGCACACATCAACAGCTGCTGGCCCCAGAAAGGCATC 3820
Qy 1048 TyrPheLysLeuValAsnAlaGlnSer 1056
Db 3821 TATTTTCCATGGTCAGTGTCCAGGCT 3847

Search completed: March 31, 2003, 02:13:48
Job time : 14525 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 30, 2003, 01:20:01 ; Search time 1034.02 Seconds
(without alignments)
2304.231 Million cell updates/sec

Title: US-09-873-409-4

Perfect score: 5330

Sequence: 1 MVISLTSKLSAYSAGAVA.....QELLNRDIYFKLVNAQSVQ 1059

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -DEV=xlh

-MODE=frame+p2n.model -DEV=xlh
-O=/cpn2.1/USPTO.spool/US09873409/runat.27032003.115416.19216/app_query.fasta_1.7544
-DB=N-Geneseq_101002 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -IOFCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=BLOSUM62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09873409 @CGN 1.1.4247 @runat.27032003.115416.19216 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DLEXT=7

Database :

N Geneseq_101002.*
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Length	DB	ID	Description
1	4985.5	93.5	3699	24	ABK83223	Human transporter
2	3115.5	58.5	4279	22	AAO03506	Dog P-glycoprotein
3	3112.5	58.4	4279	22	AAO03488	Dog P-glycoprotein
4	3112.5	58.4	4279	22	AAO03504	Dog P-glycoprotein
5	3112.5	58.4	4279	22	AAO03505	Dog P-glycoprotein
6	3111.5	58.4	4317	22	AAO03489	Dog P-glycoprotein
7	3106.5	58.3	4369	21	AAZ52047	Rat multidrug resi
8	3106.5	58.3	4425	21	AAZ52048	Rat multidrug resi
9	3100.5	58.2	3860	21	AAZ49332	Human wild-type mu
10	3100.5	58.2	3860	21	AAZ49333	Human G185V mutant
11	3100.5	58.2	3860	24	ABA94365	Human BCRP DNA rel
12	3100.5	58.2	3860	24	ABA94366	Human BCRP DNA rel
13	3100.5	58.2	4186	22	AAF6127	Cynomologous monke
14	3100.5	58.2	4195	22	AAF6128	Human P-gly
15	3100.5	58.2	4264	19	AAV66533	Mutated human P-gl
16	3100.5	58.2	4264	19	AAV66534	Human intestine ce
17	3100.5	58.2	4349	22	AAH57442	Human ATP binding
18	3100.5	58.2	4646	21	AAZ94738	Human mdri gene.
19	3100.5	58.2	4646	24	AAZ38994	Kidney cancer rela
20	3100.5	58.2	4646	24	ABL68592	Kidney cancer rela
21	3100.5	58.2	4646	24	ABL68880	Sequence of human
22	3100.5	58.2	4669	8	AAH70752	Hybrid vector pSF-
23	3100.5	58.2	6505	17	AAT13394	Retroviral M4 mar-
24	3100.5	58.2	8630	21	AAZ24041	Retroviral vector
25	3100.5	58.2	8630	21	AAZ24042	Sequence of human
26	3097.5	58.1	4669	14	AAQ52726	Mouse BCRP DNA rel
27	3090.5	58.0	4788	21	AAZ49335	Mouse BCRP DNA rel
28	3090.5	58.0	4788	21	ABA94368	Human multidrug r
29	3089.5	58.0	4646	15	AAQ72872	Human multidrug re
30	3086	57.9	4233	21	AAZ90198	Rat mdrlb2 (multis
31	3086	57.9	4233	21	AAZ27498	Rat mdrlb2 multidr
32	3085.5	57.9	4378	11	AAQ04522	Multidrug Resistan
33	3085	57.9	3924	21	AAZ94742	Human ATP binding
34	3085	57.9	3924	21	AAZ88974	Human MDR-3 DNA.
35	3085	57.9	3924	24	ABN95801	Gene #2299 used to
36	3085	57.9	4189	21	AAZ49334	Mouse multidrug r
37	3085	57.9	4189	24	ABA94367	Mouse BCRP DNA rel
38	3085	57.9	4313	14	AAQ38950	Mouse multidrug re
39	3084	57.9	3840	24	ABL91687	Human polynucleoti
40	3084	57.9	3988	21	AAZ88973	Human MDR-1 DNA.
41	3073.5	57.7	4669	19	AAV32645	Human P glycoprote
42	3073.5	57.7	4669	24	ABK52041	cDNA encoding huma
43	3044.5	57.1	4254	24	ABK63517	Rat sequence diffe
44	3043	57.1	3912	24	ABK63653	Rat sequence diffe
45	2515.5	47.2	4776	21	AAZ94744	Human ATP binding

ALIGNMENTS

RESULT 1

ABK83223

ID ABK83223 standard; cDNA; 3699 BP.

XX ABK83223;

AC ABK83223;

XX 27-AUG-2002 (first entry)

DT Human transporter and ion channel, TRICH14, Incyte ID 7472030CB1, cDNA.

DE Human; ss; gene; transporter and ion channel; TRICH; transport disorder;
XX neurological disorder; muscle disorder; immunological disorder; cancer;
KW scleroderma; systemic lupus erythematosus; allergy; leukaemia;
KW cell proliferative disorder; cervical cancer; breast cancer;
KW neurodegenerative disorder; Parkinson's disease; Alzheimer's disease;
KW myotonic dystrophy; catatonia; endocrine disorder; diabetes;
KW Grave's disease; gastrointestinal disorder; Crohn's disease;
KW renal disorder; Good pasture's syndrome; viral infection; cirrhosis;

KW bacterial infection; fungal infection; parasitic infection;
 KW protozoal infection; helminthic infection; cardiovascular disorder;
 KW atherosclerosis; hepatic disease.

XX Homo sapiens.

XX WO200240541-A2.

XX 23-MAY-2002.

XX 25-OCT-2001; 2001WO-US46055.

XX 27-OCT-2000; 2000US-243989P.

XX 03-NOV-2000; 2000US-245904P.

XX 09-NOV-2000; 2000US-247673P.

XX 17-NOV-2000; 2000US-249661P.

XX 20-NOV-2000; 2000US-252232P.

XX 01-DEC-2000; 2000US-250790P.

XX (INCY-) INCYTE GENOMICS INC.

XX Tang YT, Yue H, Nguyen DB, Hafalia AJA, Elliott VS, Lu Y;

XX Walla NK, Yao MG, Baughn MR, Gandhi AR, Ding L, Sanjanwala M;

XX Ramkumar J, Arvizu C, Gietzen KJ, Lal PG, Azimzai Y, Khan FA;

XX Thangavelu K, Thornton M, Lu DAM, Tribouley CM, Warren BA;

XX Ison CH, Das D, Raumann BE, Policky JL, Kearney L;

XX WPI; 2002-463570/49.

XX P-PSDB; ABG61544.

XX New transporters and ion channels (TRICH) polypeptides, useful for

XX diagnosing, preventing, and treating disorders associated with an

XX abnormal expression or activity of TRICH, e.g. immunological, muscular

XX or renal disorders

XX Claim 5; Page 172; 178pp; English.

XX The invention relates to human transporters and ion channels (TRICH)
 CC polypeptides, a naturally occurring amino acid sequence 90 % identical to
 CC TRICH, a biologically active fragment of TRICH or an immunogenic fragment
 CC of TRICH. Also included are an isolated polynucleotide encoding TRICH,
 CC a recombinant polynucleotide comprising a promoter sequence operably
 CC linked to the TRICH polynucleotide, a cell transformed with the
 CC recombinant polynucleotide, a transgenic organism comprising the
 CC recombinant polynucleotide, an isolated antibody that binds specifically
 CC to TRICH, and screening for compounds which bind to TRICH, modulate
 CC TRICH, modulate TRICH expression or are ant/agonists of TRICH.

CC The polypeptides are useful for diagnosing, treating, and
 CC preventing transport, neurological, muscle, immunological disorders
 CC (e.g. scleroderma, systemic lupus erythematosus, allergies), cell
 CC proliferative disorders such as cancers (e.g. leukemia, cervical or
 CC breast cancers), neurodegenerative disorders (e.g. Parkinson's disease,
 CC Alzheimer's disease), muscular disorders (e.g. myotonic dystrophy,
 CC catatonias), endocrine disorders (e.g. diabetes, Grave's disease),
 CC gastrointestinal disorders (e.g. Crohn's disease), renal disorders
 CC (e.g. Good pasture's syndrome), viral, bacterial, fungal, parasitic,
 CC protozoal and helminthic infections, cardiovascular disorders (e.g.
 CC atherosclerosis), or hepatic diseases (e.g. cirrhosis) and many
 CC other diseases and disorders detailed in the specification. They can also
 CC be used in assessing the effects of exogenous compounds on the
 CC expression of nucleic acid and amino acid sequences of transporters and
 CC ion channels. TRICH or its fragments may also be used in screening for
 CC compounds that specifically bind to and modulate the activity of TRICH.
 CC The polynucleotides can be used to create knock-in humanised animals or
 CC transgenic animals to model human disease. The present sequence
 CC encodes a TRICH protein.

XX Sequence 3699 BP; 1116 A; 707 C; 860 G; 1016 T; 0 other;

XX Alignment Scores:

Pred. No.: 0 Length: 3699

Score: 4985.50 Matches: 1002

Percent Similarity: 94.80% Conservative: 1

Best Local Similarity: 94.71% Mismatches: 2
 Query Match: 93.54% Indels: 53
 DB: 24 Gaps: 2

US-09-873-409-4 (1-1058) x ABK83223 (1-3699)

QY 1 MetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAla 20
 DB 682 ATGGTCATCTCATTGACCAAGTAAGGAATTAAGTGCCTATTCCAAAGCTGGGCTGGCA 741
 QY 21 GluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeu 40
 DB 742 GAAGAAAGTCTTGTCAATCCGAACAGTCATAGGCTTTAGGGCCCGAGAGAAAGAACTT 801
 QY 41 GlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeu 60
 DB 802 CAA----- 804
 QY 61 LeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPhe 80
 DB 805 -----AGGTATACACAGAAATCTCAAGATGCAAGAGGATTTT 840
 QY 81 GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsn 100
 DB 841 GGCATAAAAGGACTATAGCTTCAAAAGTGTCTCTTGGTGCTGTGTACTTCTTTATGAAT 900
 QY 101 GlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGly 120
 DB 901 GGNACCTATGGACTTGTCTTTTGGTATGGAACCTCTTGATCTTAATGGAGAACCTGGA 960
 QY 121 TyrThrIleGlyThrValLeuAlaValPheSerValIleHisSerSerTyrCysIle 140
 DB 961 TATACCATCGGACTCTTCTGCTGTTTCTTAGTGTAAATCCATAGTAGTATTATTCATT 1020
 QY 141 GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 160
 DB 1021 GGAGCAGCAGTCCCTCACTTTGAACCTTCGCAATAGCCGAGGAGCTGCTTTCATATT 1080
 QY 161 PheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysPro 180
 DB 1081 TTCCAGGTTATTGATGAAGAAACCCAGTATAGGTAACTTTTCCACAGCTGGATATAACCT 1140
 QY 181 GluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgPro 200
 DB 1141 GAATCCATAGAGGAACCTGTGGAATTTAAAAATGTTTCTTCAATTTATCCATCAAGGCCA 1200
 QY 201 SerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu 220
 DB 1201 TCTATCAAGATTCTGAAGGCTCTGAAATCTCGGAATTAAGTCTGGAGAGACAGTCGCCCTTG 1260
 QY 221 ValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAsp 240
 DB 1261 GTGGTCTCAATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGAT 1320
 QY 241 ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis 260
 DB 1321 CCGGATGATGGCTTTATCATGGTGGATGAGATGACATCAGAGCTTTAAATGTGCGGCAT 1380
 QY 261 TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSer 280
 DB 1381 TATCGAGACCATATTGGAGTGGTGTAGTCAAGAGCCCTGTTTGTTCGGGACCACCATCAGT 1440
 QY 281 AsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArg 300
 DB 1441 AACATATCAAGTATGGACGAGATGATGTGACTGATGAGAGATGGAGAGAGCAGCAAGG 1500
 QY 301 GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly 320
 DB 1501 GAAGCAAAATGCGTATGATTTTATCATGGAGTTCCTTAATAAATTAATATCATGCTAGGG 1560
 QY 321 GluLysGlyAlaGlnMetSerGlyGlyGlnLysGluArgIleAlaIleAlaArgAlaLeu 340
 DB 1561 GAAAAAGGAGCTCAAAATGATGGAGGGCAGAAACAGAGGATCGCAATTGCTCTGTGCTTAA 1620

QY 341 ValArgAsnProlysisLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 360
DB 1621 GTTCGAAACCCCAAGATTCTGATTTTAGATGAGGCTACGCTCCCTGGATTTCAGAAAGC 1680
QY 361 LysSerAlaValGlnAlaLeuGluLysAlaSerLysGlyArgThrThrIleValVal 380
DB 1681 AAGTCAGCTGTTCAAGCTGCACCTGGAGAAAGCGAGCAAAAGGTCCGACTCAATCGTGGTA 1740
QY 381 AlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMet 400
DB 1741 GCACACCGACTTCTACTATTTCGAGTGCAGATTGATTGTGACCCCTAAAGGATGGAATG 1800
QY 401 LeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrThrSerLeu 420
DB 1801 CTGGCGAGAAAGGAGCGCATGCTGAACATAATGGCAAAACGAGTCTATATTATTCACATT 1860
QY 421 ValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrThrSerThr 440
DB 1861 GTGATGTCAAGGATATTAAAAAAGCTGATGAACAGATGGAGTCAATGACATATTTCTACT 1920
QY 441 GluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle 460
DB 1921 GAAAGAAAGCCAACTCACTTCTCTGCACTCTGTGAGAGCATCAAGTCAGACTTCATT 1980
QY 461 AspLysAlaGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeu 480
DB 1981 GACAAGGCTGAGGAATCCACCAATCTAAAGAGATAAGTCTCTCTGAAAGTCTCTCTATTA 2040
QY 481 LysIleLeuLysLeuAsnLysProGluTyrProPheValValLeuGlyThrLeuAlaSer 500
DB 2041 AAAATTTAAAGTTAAACAAGCCCTGAATGCCCTTTTGTGGTCTTGGGGACATTTGGCTTCT 2100
QY 501 ValLeuAsnGlyThrValHisProValPheSerIleThrPheAlaLysIleThrMet 520
DB 2101 GTTCTAAATGGAACTGTTCATCCAGTATTTTCCATCATCTTTGCCAAAAATTAAACCATG 2160
QY 521 PheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePhe 540
DB 2161 TTTGGAAATAATGATAAAACCAATTAAGCATGATGCAGAAATTTATTCATGATATTC 2220
QY 541 ValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArg 560
DB 2221 GTCATTTGGGTGTATTTCCTTTGTTCAGTTATTTTCATG----- 2259
QY 561 AlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGln 580
DB 2260 -----CAG 2262
QY 581 AspIleAlaThrPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAla 600
DB 2263 GATATTGCCTGGTTTGTATGAAAGGAAAAACAGCACAGAGGCTTGACAAATATTAGCC 2322
QY 601 IleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsn 620
DB 2323 ATAGATATAGCAAAATTCAGGAGCAACAGGTTCCAGGATGGCGTCTTTAAACACAAAAT 2382
QY 621 AlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyrGluMetThrPhe 640
DB 2383 GCAACTTAACATGGGACTTTTCAGTTATCATTTCTTTATATATGATGGAGATGACATTC 2442
QY 641 LeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMet 660
DB 2443 CTGATTTCTGAGTATTGCTCCAGTACTTTCGCTGACAGGAATGATTGAAACCGGAGCATG 2502
QY 661 ThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGlu 680
DB 2503 ACTGGATTTGCCAACAAAGATAAGCAAGAACTTAAGCATGCTGGAAAGATAGCAACTGAA 2562
QY 681 AlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMet 700
DB 2563 GCTTTGGAGAAATATACGTACTATAGTGTCTATTAAACAGGGGAAAAAGCCTTCGAGCAATG 2622

RESULT 2

QY 701 TyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGly 720
DB 2623 TATGAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCACAGATTATTGGA 2682
QY 721 SerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPhe 740
DB 2683 AGCTGTATTGCAATTCAGCCATGCCITTTATATATTTTGCCTATCGCGCAGGTTTCGATTT 2742
QY 741 GlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThr 760
DB 2743 GGAGCCATTATTAATCAAGCTGGACGAATGACCCACAGAGGGCATGTTTCATAGTTTACT 2802
QY 761 AlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSer 780
DB 2803 GCAATTTGCATATGGAGCTATGCCATCGGAGAAACGCTCGTTTGGCTCCTGTAATATTC 2862
QY 781 LysAlaLysSerGlyValAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAsp 800
DB 2863 AAAGCCAAATCGGGGCTCGGCATCTGTTTGGCTTGTGGAAAAAGAAACCAATATAGAC 2922
QY 801 SerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGlu 820
DB 2923 AGCCGCAAGTCAAGAGGGAAAAAGCCACAGACACATGTGAAGGGAATTTAGAGTTTCAGAA 2982
QY 821 ValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSer 840
DB 2983 GTCTCTTTCTCTATCCATGTCGCCCAGATGTTTTCATCTCCGTGGCTTATCCCTCAGT 3042
QY 841 IleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSer 860
DB 3043 ATTGACGAGGAAAGACAGTAGCATTTGTGGGAGCAGCGGCTGTGGGAAAAACACTTCT 3102
QY 861 ValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyVal 880
DB 3103 GTTCACTTCTCGAGAGACTTTATGACCCCGTCGACAGGACAGTGTGTTGATGTGTG 3162
QY 881 AspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGlu 900
DB 3163 GATGCAAAAGAAATTGAATGTACAGTGGCTCCGTTCCAAATAGCAATCGTTCCTCAAGAG 3222
QY 901 ProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgVal 920
DB 3223 CCGTGTCTCTTCAACTGCGAGCATTTGTGAGAACATCGCTATGTCGACACACCCGTGTG 3282
QY 921 ValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGlu 940
DB 3283 GTGCCATTAGATGAGATCAAGAGAGCCGCAATGACAGCAATATCCATTCTTTATTGAA 3342
QY 941 GlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGly 960
DB 3343 GGTCTCCCTGAGAAATACAAACACACAAAGTTGGACTGAAAGGAGCACAGCTTTCTGGCGGC 3402
QY 961 GlnLysGlnArgLeuAlaIleAlaArgAlaLeuGlnLysProLysIleLeuLeuLeu 980
DB 3403 CAGAAACAAAGACTAGCTATTGCAAGGGCTCTTCTCCAAAAACCCAAAAATTTTATTGTTG 3462
QY 981 AspGluAlaThrSerAlaLeuAspAsnSerGluLysValValGlnHisAlaLeuAsp 1000
DB 3463 GATGAGCCACTTCAGCCCTCGATATGACAGTGAAGAGGTGTTTCAGCATGCCCTTGAT 3522
QY 1001 LysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsn 1020
DB 3523 AAAGCCAGGAGCGGAAAGGACATGCTGTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAAC 3582
QY 1021 AlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlu 1040
DB 3583 GCAGATTGTATAGTGTCTTCGCAATTGGAAAGATAAAGGAAACAAAGGAACTCATCAAGAG 3642
QY 1041 LeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 1058
DB 3643 CTCTTGAGAAATCGACATATATTTTAAAGTTAGTGAATGACAGTCAGTCAG 3696


```
QY 364 ValGlnAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383
Db 1727 GTTCAGTGCGCCCTGGATAAGGCAGAAAAAGCGGACCTACCATTTGTGATAGCTCATCGT 1786
QY 384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403
Db 1787 TTCTCTACACAGTTCGTAATGCCGATGTCATTTGCTGGTTTGTGATGAGTGTGGAG 1846
QY 404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423
Db 1847 AAGGAAATCATGATGAACTCATGAAGAGAGAGGCAATTTACTTCAAACTGTGCACAAATG 1906
QY 424 GlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLys 443
Db 1907 CAG---ACAAGAGGAAATGAATTTAGTTAGAAAATGCCACTGGTGAATCCAAAAGTGAA 1963
QY 444 ThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle----- 460
Db 1964 AGTGATGCCCTTGGAAATG---TCTCCAAAAGATTTCAGGGTCCAGTTTAAATAAAAAGAA 2020
QY 461 -----AspLysAlaGluGluSerThrGln 468
Db 2021 TCAACTCGCAGGAGTATACATGACCAACAGGCCAAGCAAGAAAGCTGTGTACAAAGAG 2080
QY 469 SerLysGluIleSerLeuProGluValSerLeuLysIleLeuLysLeuAsnLysPro 488
Db 2081 GACTTGAATGAGAATGTACCTCCAGTTTCCTCTGGAGGATTCGAAAGTGAACACTCAACT 2140
QY 489 GluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPro 508
Db 2141 GAATGGCCTTATTTTGTGGTGTGTATATTTGTGTATATATAAACGAGGAGCCTCCAAACCA 2200
QY 509 ValPheSerIleIlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAspLysThr 527
Db 2201 GCATTTTCAATAATATTTTCAAGGATTTATAGGATCTTTACCCGAGATGAGGATCCTGAA 2260
QY 528 ThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCys 547
Db 2261 ACAAAACGACAGAAATAGTAACATGTTTCTGTATTTCTGTCTGTCTTGAATATTTCT 2320
QY 548 PheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMet 567
Db 2321 TTTATTACATTTTCTCCAGGGCTTCACATTTGGCAAGCTTGGGAGATCTCCACTAAG 2380
QY 568 ArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGlu 587
Db 2381 CGGCTTCGATACATGTTTTCAGATCCATCTGAGCAGGATGTCAGCTGGTTGTATGAC 2440
QY 588 LysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGln 607
Db 2441 CCTAAAACACCACTGGAGCATTTGACAAACAGGCTTGCCCAATGATCGGCTCAAGTTAA 2500
QY 608 GlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSer 627
Db 2501 GGGCTATAGGTTCCAGGCTTGCTGTCATTACCCAGAATATAGCAAAATCTTGGACAGGC 2560
QY 628 ValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaPro 647
Db 2561 ATTATTATATCTTAAATCTATGTTGGCAATTAACATTTTACTCTTAGCAATTTGACCC 2620
QY 648 ValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAsp 667
Db 2621 ATCATTCGAATAGCAGAGGTTGTTGAATGAATGTTGTCTGGCAAGCACTGGAAGAT 2680
QY 668 LysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThr 687
Db 2681 AAGAAAGAGCTAGAGGAGCTGGGAAGATGCTACAGAAGCCATCGAAAACCTTCCGAAC 2740
QY 688 IleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThr 707
Db 2741 GTTGTGTTCTTGTGCTCGGAGCAGAAAGTTTGAATACATGTATGCACAGAGTTTTCGAAGTA 2800
```

```
QY 708 GlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHis 727
Db 2801 CCATACAGAAACTCTTTGAGGAAAGCACACTTTCGGGCTCTCATTTTCTACCCAG 2860
QY 728 AlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyValaTyrLeuIleGlnAla 747
Db 2861 GCAATGATGATATTTTCTTATCTGCTGCTGCTTTCCGGTTTGGTGCTTACTTGGTGGCAAT 2920
QY 748 GlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMet 767
Db 2921 GAGTTTCATGAACCTTTCAGGATGTTCTTTTGGTATTCTCAGCTATTGCTTTGGTGCCATG 2980
QY 768 AlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAla 787
Db 2981 GCAGTGGGCGAGTTCAGTTTCATTTGCTCTGATATGCCCCAAAGCCAAAGTATCAGCAGCC 3040
QY 788 HisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLys 807
Db 3041 CAGCTCATCATGATCATTTGAAAAGAGCCCTCTGATTGACAGCTACAGCCCTCAGGCGCTC 3100
QY 808 LysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCys 827
Db 3101 AAGCCAAATACGTTGGAAGAAATGTGACATTTAATGAGTCTGTTCAACTATCCCACT 3160
QY 828 ArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrVal 847
Db 3161 CGACCACAGACATCCCCGTGCTCCAGGGCTGAGCCTCGAGTGAAGAGCCAGCCAGCGCTG 3220
QY 848 AlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeu 867
Db 3221 GGCCTCGTAGGTAGCAGTGGCTGTGGGAAGAGCACAGTTGTTGAGTCTCTAGAGCGCTTC 3280
QY 868 TyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnVal 887
Db 3281 TATGACCCCTTGGCTGGTTCAGTGCTAATGTATGTCGCAAGAGATAAAGCACCTGATGTC 3340
QY 888 GlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSer 907
Db 3341 CAGTGGCTCCGAGCACACCTGGGCATCGTGTCTCAGAGAGCCCATCTCTGTGTGACTGCAGC 3400
QY 908 IleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLys 927
Db 3401 ATTGCCGAGAAATTCCTTATGGAGACAAACAGCCGGTCTGATFCAATGAAGAGATTATG 3460
QY 928 GluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsn 947
Db 3461 CAGCAGCCCAAGAGGAGCCAAACATACACCACTTCATCGACACACTCCCTCGAATAATACAAC 3520
QY 948 ThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIle 967
Db 3521 ACCAGAGTAGGAGACAAAGGAACCCAGCTCTCTGTGGCCAGAAACAGGCGCATTTGCCATA 3580
QY 968 AlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeu 987
Db 3581 GCTCGCGCTCTTGTATAGACAGCCTCATATTTGCTTTTGGATGAAGCTACATCAGCTCTG 3640
QY 988 AspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThr 1007
Db 3641 GATACAGAAAGTGAAAGGTTGTCCAAAGAGCCCTCGACAAAGCCAGAGAGAGCCGCCACC 3700
QY 1008 CysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeu 1027
Db 3701 TGCATTGTGATCGCCCAACCGCTTGTCCACCATCCAGATGTCAGATTTAATAGTGGTGT 3760
QY 1028 HisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIle 1047
Db 3761 CAGAATGGCAAAAGTCAAGGAGCATGGCACACATCAACAGCTGCTGCTGCCCAAGAGCATC 3820
QY 1048 TyrPheLysLeuValAsnAlaGlnSer 1056
Db 3821 TATTTTCCATGTCAGTGTCCAGGCT 3847
```

RESULT 4

AAD03504
ID AAD03504 standard; cDNA; 4279 BP.

AC AAD03504;

XX 13-JUN-2001 (first entry)

XX Dog P-glycoprotein (PGP) allelic variant (Genotype A) cDNA.

KW Dog; P-glycoprotein allelic variant; multidrug transporter; MDR1;
KW drug bioavailability; transgenic animal; genetic model; ss.

XX Canis familiaris.

XX Key Location/Qualifiers

FT 17..3862

FT /tag= a

FT /product= "Dog P-glycoprotein (PGP) allelic variant

FT (Genotype A) protein"

FT replace (607, C)

FT /*tag= b

XX WO200123540-A2.

XX 05-APR-2001.

XX 28-SEP-2000; 2000WO-US26767.

XX 28-SEP-1999; 99US-0156510.

XX (GENT-) GENTEST CORP.

XX Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;

XX WPI; 2001-235373/24.

XX P-PSDB; AAE00308.

XX New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
XX for determining the bioavailability of drugs and for screening for dog
XX PGP inhibitors -

XX Claim 9; Page 85-90; 11pp; English.

XX The invention relates to dog P-glycoprotein (PGP) also referred
XX as multidrug transporter (MDR1) and nucleic acids encoding them.
XX The invention also includes fragments and biologically functional
XX variants of dog P-glycoprotein. PGP and their nucleic acids are
XX useful for determining the bioavailability of drugs and for
XX screening PGP inhibitors. They are useful for the diagnosis and
XX treatment of conditions characterised by PGP activity, by
XX reducing or increasing PGP activity in a cell. PGP nucleic acids
XX are used as oligonucleotide probes. Complements of PGP nucleic
XX acids are useful as antisense oligonucleotides, to induce a PGP
XX 'knockout' phenotype. They are used to prepare a non-human
XX transgenic animal, which are valuable as genetic models for
XX human diseases.

XX The present sequence is dog P-glycoprotein (PGP) allelic variant
XX (Genotype A) cDNA. The PGP enzyme functions as an efflux pump
XX exporting small molecules across the cell membrane. This enzyme
XX is a member of the ABC transporter family.

XX Sequence 4279 BP; 1295 A; 833 C; 1008 G; 1143 T; 0 other;

XX Alignment Scores:

Pred. No.:	1.61e-283	Length:	4279
Score:	3112.50	Matches:	601
Percent Similarity:	75.02%	Conservative:	201
Best Local Similarity:	56.22%	Mismatches:	220
Query Match:	58.40%	Indels:	47
DB:	22	Gaps:	6

XX US-09-873-409-4 (1-1058) x AAD03504 (1-4279)

QY	4	SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluVal	23
Db	734	TCATTTACTGATAAAGAACTCTTGGCTATGCAAAAGCTGAGCAGTAGCTGAAGAATC	793
QY	24	LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer	43
Db	794	TTAGCGCAATCAGAACTGTGATTGCTTTGGAGGACAAAGAAAGAACTTGA-----	847
QY	44	PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys	63
Db	847	-----	847
QY	64	ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys	83
Db	848	-----AGGTACAACCAAAATTTAGAGAAGCTAAAGAAATTTGGATAAAG	892
QY	84	ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr	103
Db	893	AAAGCTATCACGGCCCAACTTCTATTGTGGCGCTTTCTTATTGATCTATGTCATCAT	952
QY	104	GlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle	123
Db	953	GCTCTGGCTTTCTGGTATGGACCTCTTGGTCTCTCCAGTGAA-----TATTCTATT	1006
QY	124	GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla	143
Db	1007	GGACAAGTACTCACTGTCTTTCTTCTATTAAATTTGGGGCTTTTAGTATTGGACAGGA	1066
QY	144	ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal	153
Db	1067	TCCCCAAGCATTTGAAGCATTTGCCAAACGCAAGAGGAGCTTATGAAATCTTCAAGATA	1126
QY	164	IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle	183
Db	1127	ATTGACAATAAACCAAGCAATTGACACTATTGGAAGCTGGACATTAACACGATATATT	1186
QY	184	GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys	203
Db	1187	AAGGGAAATTTGGAAATTTCAAAATTTTCACTTCACTTCACTTCTCGAAAAGATTAAG	1246
QY	204	IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu	223
Db	1247	ATCTTAAAGGCTCTCAACCTGAAGTTTCAGAGTGGCAGCAGCTGGCTGGTTGGAAAC	1306
QY	224	AsnGlySerGlyLysSerThrValValGlnLeuGlnArgLeuTyrAspProAspAsp	243
Db	1307	AGTGGCTGGGGAGAGACGACCGTGCAGCTGATGCAGAGGCTCTATGACCCACAGAT	1366
QY	244	GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp	263
Db	1367	GGCATGGTCTGTATTGATGGACAGGACATTAGGACCAATAAATGTAAGGCATCTTCGGAA	1426
QY	264	HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnIle	283
Db	1427	ATTACTGGTGTGGTGGTGGAGGCTGTGTGTGTTCACCCACCATAGCTGAAACATT	1486
QY	284	LysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn	303
Db	1487	CGCTATGCCCGCAAAATGTCACCATGGATGAGATGAGAAAGCTGTTAAGGAAGCCAT	1546
QY	304	AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyLysGly	323
Db	1547	GCCTATGATTTTATCATGAAACTACCTAATAAATTTGACACTCTGGTTGGAGAGAGGG	1606
QY	324	AlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn	343
Db	1607	GCCAGCTGAGTGGTGGACAGAAACAGAGAAATCGCAATGCTCGGGCCCTGGTTGCAAC	1666
QY	344	ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla	363
Db	1667	CCCAAGATTTCTTCTGCTGATGAGGCAACGCTCAGCTCTGGACACTGAAAGTGAAGCAGT	1726
QY	364	ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg	383

	Db	1727	GTTCAAGTGGCCCTGATTAAGCGCAGAAAGACCGGACTCACATTGTGATAGTCATCATCGT	1786
	Qy	384	LeuSerThrIleArgSerAlaAspLeulleValThrLeuLysAspGlyMetLeuAiaGlu	403
	Db	1787	TTCGTCTACAGTTCTGAATGCCGATGCATTCGTGTTTTGATGATGANGTCATGTGGAG	1846
	Qy	404	LysGlyAlaHisGlaGluLeuMetAlaLysArgGlyLeuTytyrSerLeuValMetSer	423
	Db	1847	AAAGGAAATCATGATGAATCTATGAAGAAGAGAGGCCATTTACTTTCAAATTTGTCAACAATG	1906
	Qy	424	GlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrtYrSerThrGluArgLys	443
	Db	1907	CAG---ACAAGAGGAAATGAATAATGAGTTAGAAAATGCCACTGGTGTAATCCAAAAGTAA	1963
	Qy	444	ThrAsnSerLeuProLeuHisSerVallysSerIleLysSerAspPheIle-----	460
	Db	1964	AGTGATGCCTTGGAAATG---TCTCCAAGAAGNATCAGGGTCCAGTTTAATAAAAAAGAGA	2020
	Qy	461	-----AspLysAlaGluGluSerThrGln	468
	Db	2021	TCAACTCGCAGGAGTAGTATACATGCACCACAAGGCCAAGACAGAAAGCTTGGTACAAAAGAG	2080
	Qy	469	SerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysPro	488
	Db	2081	GACTTGAATCAGAAATGACTCTCCAGTTTCCTTCTCGGAGGATTCGAAAGCTGAATCAACT	2140
	Qy	489	GluTrpPropheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPro	508
	Db	2141	GAATGCCCTATTTCGTGGTGGTATATTTGCTGTTATTAAACGGAGCGCTGCACACCA	2200
	Qy	509	ValPheSerIleIlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAspLysThr	527
	Db	2201	GCATTTTCAATATATTTTCAAGGATTAAGGATCTTTACCAGAGTAGGATCCTGAA	2260
	Qy	528	ThrLeuLysHisAspAlaGluIleTyrsSerMetIlePheValIleLeuGlyValIleCys	547
	Db	2261	ACAAAACGACAGAATAGTAACATGTTTCTGTATTGTTCTAGTCCCTGGAAATATTATTCT	2320
	Qy	548	PheValSerTyPheMetGlnGlyLeuPheTyrglyArgAlaGlyGluIleLeuThrMet	567
	Db	2321	TTTATTACATTTTCTCCAGGCTTCACATTTGGCAAGCTGGGAGATCCTCACTAAG	2380
	Qy	568	ArgLeuArgHisLeuAlaPheLysAlaMetLeutyrginAspIleAlaTtpPheAspGlu	587
	Db	2381	CGCTTCGATACATGCTTTTCAGATCCATCTGTGACACAGGATGCAGCTGGTTTGATGC	2440
	Qy	588	LysGluAsnSerThrGlyGlyLeuthrThrIleLeuAlaIleAspIleAlaGlnIleGln	607
	Db	2441	CCTAAAAACACCACTCGAGCATTGCAACACGAGCTTGCCAATGATGGGCTCAAGTTAAA	2500
	Qy	608	GlyAlaThrGlySerArgileGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSer	627
	Db	2501	GGGGCTATAGTTCCAGGCTTCTGCTATTTCCCAGATATAGCAATCTTGGGACAGGC	2560
	Qy	628	ValIleIleSerPheIleTyrglyTrpGluMetThrPheLeuilleLeuSerIleAlapro	647
	Db	2561	ATTATTATATCTTATCTATGTTGGCAATTAACACTTTTACTCTTAGCAATGTGATACC	2620
	Qy	648	ValLeuAlavalThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAsp	667
	Db	2621	ATCATTTGCAATAGCAGGATTTGTGAAATGAAAAATGTTGTTCTGGCAGACGACTGAAAGAT	2680
	Qy	668	LysGluCluleuLysHisalaGlyLysIleAlaThrGluAlaLeuGluAsnIleArqThr	687
	Db	2681	ARGAAAGAGCTAGAGGAGCTGGGAAGATTGTCACAGAAGCCATCGAAAAATCTCCGAAC	2740
	Qy	688	IleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrgluGluMetLeuGlnThr	707
	Db	2741	GTGTGTTCTTGTCTCGGAGCAGAAAGTTTGAATACATGATGCACAGAGTTTGCAGATA	2800
	Qy	708	GlnHisArgAsnThrSerLysLysAlaGlnIlelleGlySerCystyrAlaPheSerHis	727

Db	2801	CCATACAGAAACTCTCTTGGAGGAAAGACACATCTCTTCGGGGTCTCAATTTTCTATCACCAG	2865
Qy	728	AlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyValaTyrLeuIleGlnAla	747
Db	2861	GCAATGATGATATTTTCCATGCTGCCTGTTTCGGGTTCCTACTCTTGGTGCCAAAT	2920
Qy	748	GlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMet	767
Db	2921	GAGTTCATGAACATTTCCAGGATGTTCTTTTGGTATTTCTCAGCTATTGTCTTGTGGTGCCATG	2980
Qy	768	AlaIleGlyIysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAla	787
Db	2981	GCAGTGGGCGAGTCAGTTCATTGCTCTGACTATGCCAAGGCCAAAGTATCAGCAGCC	3040
Qy	788	HisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLys	807
Db	3041	CAGTCTCATGATCATTTGAAAAAGCCCTCTGATTGACAGCTACAGCCCTCAGGGCTC	3100
Qy	808	LysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCys	827
Db	3101	AAGCCAAATACGTTGGAAGGAAATGTGCATTTAATAGAGTCTGTCTTCAACTATCTCCACT	3160
Qy	828	ArgProAspValPheIleLeuArgGlyLeuSerLeuIleGluArgGlyLysThrVal	847
Db	3161	CGACCAACATCCCCGTCTCAGGGGCTGAGCCCTCAGAGTGAAGAGGGCCAGCGCT	3220
Qy	848	AlaPheValGlySerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeu	867
Db	3221	GCCCTCGTAGTAGCAGTGGCTGTGGAGAGACACAGTTGTTACGCTCTTAGAGGGCTC	3280
Qy	868	TyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnVal	887
Db	3281	TATGACCCCTTGGCTGGTTCAGTGTCTAATGATGGAAGAGATAAAGCACCTGAATGTC	3340
Qy	888	GlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSer	907
Db	3341	CAGTGGCTCCGAGCACACCTCGGGCATCGTGTCTCAGAGGCCCATCTCTGTTCAGTCAGC	3400
Qy	908	IleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLys	927
Db	3401	ATTGCCAGAGAACATTGCCTATGGAGACACACCCGGTCTGTATCATCATGAAGAGATTATG	3460
Qy	928	GluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsn	947
Db	3461	CAGCAGCCCAAGGAGGCCAACATACACCACTTTCATCGAGACACTCCCTGAGAAATACAA	3520
Qy	948	ThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIle	967
Db	3521	ACCAGATGGAGACAAAGAACCCAGCTCTCTGTGGCCAGAAACAGCGCATTTGCCATA	3580
Qy	968	AlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeu	987
Db	3581	GCTCGGCTCTTGTAGACAGCTCATATTTTGTCTTTGGATGAAGCTACATCAGCTCTG	3640
Qy	988	AspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThr	1007
Db	3641	GATACAGAAAGTGAAAGGTTGTCGAAGAAGCCCTGGCAAGCCAGAGAGGGCGCAC	3700
Qy	1008	CysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeu	1027
Db	3701	TGCATTGTGATCCCCACCGCTGTCTCCACCATCCAGAAATGCAGATTTAATAGTGTGTT	3760
Qy	1028	HisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIle	1047
Db	3761	CAGAAATGGCAAGTCAAGGAGCATGGCACACATCAACAGCTGCTGGCCCCAGAAAGGCATC	3820
Qy	1048	TyrPheLysLeuValAsnAlaGlnSer	1056
Db	3821	TATTTTCCATGGTCAGTCTCCAGGCT	3847
RESULT	5		
ID	AAD03505		
	AAD03505 standard; cDNA; 4279 BP		

RESULT 5
AAD03505
ID AAD

XX
AC AAD03505;
XX
XX 13-JUN-2001 (first entry)
XX
XX Dog P-glycoprotein (PGP) allelic variant (Genotype B) cDNA.
XX
XX Dog; P-glycoprotein allelic variant; PGP; multidrug transporter;
KW MDR1; drug bioavailability; transgenic animal; genetic model; ss.
XX
XX Canis familiaris.
OS
XX
XX Location/Qualifiers
FH Key 17..3862
FT CDS /tag= a
FT /product= "Dog P-glycoprotein (PGP) allelic variant
FT (Genotype B) protein"
FT allele replace (91, T)
FT /tag= b
FT allele replace (607, C)
FT /tag= c
XX
XX WO200123540-A2.
XX
XX 05-APR-2001.
XX
XX 28-SEP-2000; 2000WO-US26767.
XX
XX 28-SEP-1999; 99US-0156510.
XX
XX (GENT-) GENTEST CORP.
XX
XX Stocker PJ, Steimel-crespi DT, Crespi Ch, Reif TC, Patten CJ;
PI WPI; 2001-235373/24.
XX
XX P-PSDB; AAE00309.
XX
XX New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
PT for determining the bioavailability of drugs and for screening for dog
PT PGP inhibitors -
XX
XX Claim 9; Page 93-99; 111pp; English.
XX
XX The invention relates to dog P-glycoprotein (PGP) also referred
CC as multidrug transporter (MDR1) and nucleic acids encoding them.
CC The invention also includes fragments and biologically functional
CC variants of dog P-glycoprotein. PGP and their nucleic acids are
CC useful for determining the bioavailability of drugs and for
CC screening PGP inhibitors. They are useful for the diagnosis and
CC treatment of conditions characterised by PGP activity, by
CC reducing or increasing PGP activity in a cell. PGP nucleic acids
CC are used as oligonucleotide probes. Complements of PGP nucleic
CC acids are useful as antisense oligonucleotides, to induce a PGP
CC 'knockout' phenotype. They are used to prepare a non-human
CC transgenic animal, which are valuable as genetic models for
CC human diseases.
CC The present sequence is dog P-glycoprotein (PGP) allelic variant
CC (Genotype B) cDNA. The PGP enzyme functions as an efflux pump
CC exporting small molecules across the cell membrane. This enzyme
CC is a member of the ABC transporter family.
XX
XX Sequence 4279 BP; 1296 A; 833 C; 1008 G; 1142 T; 0 other;
SQ

Alignment Scores:
Pred. No.: 1.61e-283 Length: 4279
Score: 3112.50 Matches: 601
Percent Similarity: 75.02% Conservative: 201
Best Local Similarity: 56.22% Mismatches: 220
Query Match: 58.40% Indels: 47
DB: 22 Gaps: 6

US-09-873-409-4 (1-1058) x AAD03505 (1-4279)

QY 4 SerLeuThrSerLysGluLeuSerAlaTyrrSerLysAlaGlyAlaValAlaGluVal 23
DB 734 TCATTACTGATAAAGAACTCTGGCCATATGCAAAAGCTGGAGCTAGCTGAGAGTTC 793
QY 24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43
DB 794 TTAGCAGCAATCAGAACTGTGATTGCTTGGAGGACAAAGAAAGAACTTGAA-----847
QY 44 PheLeuLeuAsnIleThrArgTyrrAlaTrpPheTyrrPheProGlnTrpLeuLeuSerCys 63
DB 847 -----847
QY 64 ValLeu***PheValArgTyrrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
DB 848 -----AGGTACAAACAAAAATTTAGAAAGAGCTAAAGAAATTTGGATAAAG 892
QY 84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrrPhePheMetAsnGlyThrTyrr 103
DB 893 AAAGCTATCAGGCCAACATTTCTATTGTGGCGCTTCTTATTGATCTCTATGCATCATAT 952
QY 104 GlyLeuAlaPheTrpTyrrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrrThrIle 123
DB 953 GCTCTGGCTTTCTGGTATGGGACCTCTTGTCTCTCCCTCCAGTGA-----TATTCTATT 1006
QY 124 GlyThrValLeuAlaValPheSerValIleHisSerSerTyrrCysIleGlyAlaAla 143
DB 1007 GGACAAAGTACTCCTCTCTCTTTCTGTATTAAATTTGGGGCTTTTAGTATTTGGACAGGCA 1066
QY 144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163
DB 1067 TCCCAAGCAATTGAGCATTTGCCAAGCGAAGAGGAGAGCTTATGAAATCTTCAAGATA 1126
QY 164 IleAspLysGlyProSerIleAspAsnPheSerThrAlaGlyTyrrLysProGluSerIle 183
DB 1127 ATTGACAATAAACAAGCATTTGACAGCTATTGCAAGAGTGGACATAAACAGATAATATT 1186
QY 184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrrProSerArgProSerIleLys 203
DB 1187 AGGGAAATTTGGAATTTCAAAATTTCTACCTTCAGTACCTTCTCGAAAAGAAAGTAAAG 1246
QY 204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
DB 1247 AUCTTAAAGGCTCAACTGAGGTTTCAGAGTGGGACAGCAGTGGCGTGGTGGGAAC 1306
QY 224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrrAspProAspAsp 243
DB 1307 AGTGGCTGCGGGAAGAGCAGCAGCGTGAGTGTGATGATGATGATGATGATGATGATGAT 1366
QY 244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrrArgAsp 263
DB 1367 GGCATGCTCTGTATTGATGGACAGACATTTAGGACCATTAATTAAGGCATCTTCGGGAA 1426
QY 264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283
DB 1427 ATTACTGGTGTGTGAGTCAAGGAGCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1486
QY 284 LysTyrrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303
DB 1487 CGCTATGGCGCGCAAAATGTCCACCATGGATGAGATTGAGAAAGCTGTTAAGGAAGCAAT 1546
QY 304 AlaTyrrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGly 323
DB 1547 GCCTATGATTTTATCATGAAACTACCTTAATAATTTGACACTCTGGTTGGAGAGAGGG 1606
QY 324 AlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343
DB 1607 GCCAGCTGAGTGTGGACAGAAACAGAGAAATGCCATTTGCTGGGCCCTGGTTCCGCAAC 1666
QY 344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
DB 1667 CCAAGATTTCTTCTGCTGGATGAGGCAACGTCAGCTCTGGACACTGAAAGTGAAGCAGTG 1726
QY 364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383

[illegible]

RESULT 6
AAD03489
ID AAD03489 standard; CDNA: 4317 BP.

XX AAD03489;
 XX AC
 XX DT
 XX DE
 XX Dog P-glycoprotein (PGP) cDNA #2.
 XX KW
 XX Dog; P-glycoprotein; PGP; multidrug transporter; MDR1;
 XX drug bioavailability; transgenic animal; genetic model; ss.
 XX OS
 XX Canis familiaris.
 XX FH
 XX Key Location/Qualifiers
 XX CDS 70..3912
 XX /tag= a
 XX /product= "Dog P-glycoprotein (PGP) #2"
 XX PN WO200123540-A2.
 XX PD
 XX 05-APR-2001.
 XX PF 28-SEP-2000; 2000WO-US26767.
 XX PR 28-SEP-1999; 99US-0156510.
 XX PA (GENT-) GENTEST CORP.
 XX PI Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;
 XX WI; 2001-235373/24.
 XX DR P-PSDB; AAE00304.
 XX PT New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
 XX for determining the bioavailability of drugs and for screening for dog
 XX PGP inhibitors -
 XX PS Claim 1; Page 66-72; 11pp; English.
 XX CC The invention relates to dog P-glycoprotein (PGP) also referred
 XX as multidrug transporter (MDR1) and nucleic acids encoding them.
 XX The invention also includes fragments and biologically functional
 XX variants of dog P-glycoprotein. PGP and their nucleic acids are
 XX useful for determining the bioavailability of drugs and for
 XX screening PGP inhibitors. They are useful for the diagnosis and
 XX treatment of conditions characterized by PGP activity, by
 XX reducing or increasing PGP activity in a cell. PGP nucleic acids
 XX are used as oligonucleotide probes. Complements of PGP nucleic
 XX acids are useful as antisense oligonucleotides, to induce a PGP
 XX 'knockout' phenotype. They are used to prepare a non-human
 XX transgenic animal, which are valuable as genetic models for
 XX human diseases.
 XX CC The present sequence is dog P-glycoprotein (PGP) cDNA. The
 XX PGP enzyme functions as an efflux pump exporting small molecules
 XX across the cell membrane. This enzyme is a member of the ABC
 XX transporter family.
 XX SQ Sequence 4317 BP; 1293 A; 844 C; 1019 G; 1161 T; 0 other;
 Alignment Scores:
 Pred. No.: 2,03e-283 Length: 4317
 Score: 3111.50 Matches: 600
 Percent Similarity: 75.02% Conservative: 202
 Best Local Similarity: 56.13% Mismatches: 220
 Query Match: 58.38% Indels: 47
 DB: 22 Gaps: 6

US-09-873-409-4 (1-1058) x AAD03489 (1-4317)

QY 4 SerLeuThrSerLysGluLeuSerAlaThrSerLysAlaGlyAlaValAlaGluVal 23
 DB 784 TCATTACTGATTAAGAACTCTTGGCCCTATGCAAAAGCTGAGCAGTAGCTGAAGAATC 843
 QY 24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43

Db 844 TTAGCAGCAATCAGAACTGTGATTGCTTTGGAGGACAAAAGAAAGAACTTGAA----- 897
 QY 44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63
 Db 897 ----- 897
 QY 64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
 Db 898 -----AGGTACAACAATAATTAGAGAAGCTAAAGAATTGGATTAAG 942
 QY 84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
 Db 943 AAAGCTATCACGGCCAACTTCTATTGGTGGCGCTTCTTATTGATCTATGATCATAT 1002
 QY 104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123
 Db 1003 GCTCTGGCTTTCTGGTATGGACCTCTCTGGTCTCTCCAGTGAA-----TATATATT 1056
 QY 124 GlyThrValLeuAlaValPhePheSerValIleHisSerTyrCysIleGlyAlaAla 143
 Db 1057 GGACAGTACTCACTGCTCTTTCTGTATTAAATGGGGCTTTAGTATTGCACAGGCA 1116
 QY 144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163
 Db 1117 TCCCCAAGCATTTGAAGCATTTGCCAAACGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1176
 QY 164 IleAspLysIleProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
 Db 1177 ATTGACAATAAACCAAGCATTTGACAGCTATTCCGAAGTGAGCATAAACCAAGATAATATT 1236
 QY 184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
 Db 1237 AAGGGAAATTGGAATTCAAAATGTTCACTTCAGTTCACCTTCTCGAAAAGAGTTAAG 1296
 QY 204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
 Db 1297 ATCTTAAGGGTCTCAACTGAAGTTTCAGAGTGGCGAGCAGCTGGCGCTGTTGGGAAC 1356
 QY 224 AsnGlySerGlyLysSerThrValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243
 Db 1357 AGTGGCTGGGGGAAGACGACGACCGTGACGTGTATGAGAGGCTCTATGACCCACACAT 1416
 QY 244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
 Db 1417 GGATGGTCTGTATTGTCGACAGGACATTAGACCATAATGTAGGCATCTTCGGGAA 1476
 QY 264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283
 Db 1477 ATTACTGGTGTGGTGTGAGTCAGGAGCCCTGTGTTGTTGCCACACGATAGCTGAAACATT 1536
 QY 284 LysTyrGlyArgAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsn 303
 Db 1537 CGCTATGGCGCGAAATGTCACTATGATGAGATGAGAAAGCTGTTAAGGAAGCCAAAT 1596
 QY 304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323
 Db 1597 GCCTATGATTTTATCATGAAACTACCTAATAATTTGACACTCTGTTGGAGAGAGAGGG 1656
 QY 324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343
 Db 1657 GCCCGGCTGAGTGGTGACAGAAACAGAGAATCGCCATGTCTGGGCGCTTGGTTCGCAAC 1716
 QY 344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
 Db 1717 CCAAGATTCTCTGCTGGATGAGCAAGCTCAGCTCTGGACACTGAAAGTAGAGAGTGT 1776
 QY 364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383
 Db 1777 GTTCAGGTGGCCCTGGTAAGGCGCAGAAAAGGCGGACCTACCATTTGTATGATCATCGT 1836
 QY 384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403

DE Rat multidrug resistance protein 1a cDNA.

XX Multidrug resistance protein 1a; mdrla; multi-specific drug transporter;
 KW drug formulation; formulation excipient; compound design; inflammation;
 KW pharmacokinetic; oral absorption; cancer; cardiovascular disease;
 KW central nervous system disorder; auto-immune disease; kidney disease; ss.
 XX Rattus rattus.

XX Key Location/Qualifiers
 FT CDS 352..4170
 FT /tags= a
 FT /product= "Rat multidrug resistance protein"

XX W0200015650-A1.

XX 23-MAR-2000.

XX 10-SEP-1999; 99WO-US20770.

XX 17-SEP-1998; 98US-0156800.

XX 09-DEC-1998; 98US-0208809.

XX (SMIK) SMITHKLINE BEECHAM CORP.

XX Brun KA, Chenery RJ, Ellens H, Feild JA, Yue L;
 WPI; 2000-271372/23.
 DR P-PSDB; AAY70596.

XX Isolated rat mdrla polynucleotides and polypeptides, useful in assays
 PT to provide information on drug formulation, selection of formulation
 PT excipients and compound design -

XX Claim 2; Page 23-24; 33pp; English.

XX The present cDNA sequence encodes rat multidrug resistance
 CC protein 1a (mdrla). This protein is a member of the multi-specific drug
 CC transporters family. Mdrla is used in assays to provide information on
 CC drug formulation, selection of formulation excipients and compound
 CC design. They are used in cell based, membrane based, binding or other
 CC assays to provide information that may enhance drug formulation. This
 CC invention further relates to the generation of in vivo and in vitro
 CC comparison data to predict oral absorption and pharmacokinetics. This
 CC enables the selection of drugs with optimal pharmacokinetics, i.e. good
 CC oral bioavailability, brain penetration, plasma half life, and minimum
 CC drug interaction. Transgenic and knock-out animals created using rat
 CC mdrla provides an insight into treating and preventing human diseases
 CC including cancer, inflammation, cardiovascular disease, central nervous
 CC system disorders, auto-immune and kidney disease.

XX Sequence 4369 BP; 1241 A; 973 C; 1096 G; 1059 T; 0 other;

XX Alignment Scores:

Pred. No.:	6.13e-283	Length:	4369
Score:	3106.50	Matches:	600
Percent Similarity:	75.54%	Conservative:	206
Best Local Similarity:	56.23%	Mismatches:	218
Query Match:	58.28%	Indels:	43
DB:	21	Gaps:	7

US-09-873-409-4 (1-1058) x AA252047 (1-4369)

QY 4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluVal 23
 DB TCATTTTACTATAGGAACCTCCAGCTTATGCAAAAGCTGAGCAGTGTGCTGAAGAATC 1098
 QY 24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43
 DB TTAGCACCATCAGAACTGTGATTGCCTTTGGAGGACAAAGAGGAACTTGAA----- 1152
 QY 44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63

Db 1152 ----- 1152

QY 64 ValLeu**PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 1197
 DB 1153 -----AGGTACAATAACATTTTGAAGAAGCTAAAGAGCTTGGGATAAG 1197

QY 84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 1257
 DB 1198 AAAGCTATCACGGCCAAACATTTCCATGGGTGCAGCTTTTCTGCTTATCTATCATCATAT 1257

QY 104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyThrThrIle 123
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 1311
 DB 1258 GCTCTGGCATTCGGTATGGGACTTCCTTGGTCATCTCAAAAGAA-----TACACTATT 1311

QY 124 GlyThrValLeuAlaValPhePheSerValIleHisSerTyrCysIleGlyAlaAla 143
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 1371
 DB 1312 GGACAAGTCTCACTGTCCTTTTCTGTATTAAATGGAGCATTCAGTGTGGCAGGCA 1371

QY 144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 1431
 DB 1372 TCTCCAATATTGAAGCTTCGCCAATGCTAGAGGAGCAGCTTATGAAGTCTTCAAGTATA 1431

QY 164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 1491
 DB 1432 ATTGATAAATAAGCCCACTATAGACAGCTTCTCAAGAGTGGGCACAAACCCGACACATA 1491

QY 184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 1551
 DB 1492 CAAGGAATTTGGAAATTCAAAATATTCTACTTCAGTATCCGCTCTCGAAAAGACGTTCAG 1551

QY 204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 1611
 DB 1552 ATCTTGAAGGGCTCAACCTGAAGGTGAAGCGCGGACAGCGTAGCCCTGGTTGGCAAC 1611

QY 224 AsnGlySerGlyLysSerThrValValGlnLeuGlnArgLeuTyrAspProAspAsp 243
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 1671
 DB 1612 AGTGGCTGTGGGAAAAGCAACTGTCCAGCTGCTCCAGAGGCTTACGACCCCATAGAG 1671

QY 244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 1731
 DB 1672 GCGAGGTCAATGATCCAGCGACAGGACATCAGACCATCAATGTGAGGTATCTCGCGGAA 1731

QY 264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 1791
 DB 1732 ATCATTTGGGTGTGAGTCAGAAACCGTCTGTTGCCACCAACAATTCGCGAAAACATT 1791

QY 284 LysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 1851
 DB 1792 CGCTATGGCCGAGAAAACGTCACCATGGATGATAGAGAAAGCTGTCAAGGAAGCCAAT 1851

QY 304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 1911
 DB 1852 GCCTATGATTTCAATGATAAATGCCCCCAAAATTTGACACCTGTTGGTGTGAGAGGG 1911

QY 324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 1971
 DB 1912 GCGCAGCTGAGTGGGGACAGAAACAGAGGATCGCCATTCGCCGGGCCCTGGTCCGCAAC 1971

QY 344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 2031
 DB 1972 CCCAAGATCCTTTTGTGGATGAGGCCACGTCAGCTTGGACACAGAAACCGAAGCGGTG 2031

QY 364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 2091
 DB 2032 GTTCAGCCCTCTGGATAGGCTAGAGAGCGCGGACCAACCATTTGTAGTAGTCCACGC 2091

QY 384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 2151
 DB 2092 TTGTCTACAGTTCGCAATGCTGACGTCATTGCTGTTTGTGTTGCTGCTGCTGCTGCTG 2151

QY 404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 2211
 DB 2152 CAAGGAATCATGATGAGCTCATGAGAGAGAAAGGAATTTACTTCAAACTTGTGCTGACT 2211


```
Qy 424 Gln-----AspIleLysLysAlaAspGluGlnMetGluSer----- 435
|||
Db 2212 CAGACAGCAGGAAATGAAATTCGAATAGGAATGAAGCTTGTGTAATCTAAAGACGGAATT 2271
|||
Qy 436 -----MetThrTySerThrGluArgLysThrAsnSerLeu-----ProLeuHisSer 451
|||
Db 2272 GATAATGTGCAGATGCTCTCAAAAGATTCGGGATCCAGTCTAATAAGAAAGAGATCAACT 2331
|||
Qy 452 ValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThr---GlnSerLys 470
|||
Db 2332 CGCAAAAGCATCGTGGGCCACATGATCAAGACGGGAACTTAGCACCAAGAGGCTCTG 2391
|||
Qy 471 GlnIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrp 490
|||
Db 2392 GATGACGACGTACCTCCAGCTTCCTTTGGCGGATCCTGAAGTTGAATTCAACTGAATGG 2451
|||
Qy 491 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 510
|||
Db 2452 CCTTATTTTGTGGTGTATTTTGTGCATATAAATGGAGGCTTGCACCCAGCATTC 2511
|||
Qy 511 SerIlePheAlaLysIleThrMetPheGlyAsnAsnAspLysThrThrLeu--- 529
|||
Db 2512 TCATATATATTTTCAAGGTGTAGGGGTTTTTACAAAATAAGACACCCCTGAAATCCAG 2571
|||
Qy 530 LysHisAspAlaGluIleTySerMetIlePheValIleLeuGlyValIleCysPheVal 549
|||
Db 2572 CGGCAGAACAGCAACTGTTTCTTATTTGTTCTGATCCTGGGATCATCTCTTTCAAT 2631
|||
Qy 550 SerTyPheMetGlnGlyLeuPheTyArgAlaGlyGluIleLeuThrMetArgLeu 569
|||
Db 2632 ACGTTTTTCTTCAGGGCTTCACATTTGGCAAGCTGGAGAGATCCTCACCAGCGACTC 2691
|||
Qy 570 ArgHisLeuAlaPheLysAlaMetLeuTyArgAlaGlyGluIleLeuThrMetArgLeu 589
|||
Db 2692 CGATACATGCTTCAATATCCATGCTGAGACAGACATAGCTGGTTGATGACCCCTAAA 2751
|||
Qy 590 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 609
|||
Db 2752 AACACACAGGAGCGCTGACCACGAGCTTGCCTCAATGAGCTCTCAAGTGAAGGGCT 2811
|||
Qy 610 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629
|||
Db 2812 ACAGGCTTAGCTGCTGTTATTTACCCAGACATAGCAATCTTGGGACAGCATCATC 2871
|||
Qy 630 IleSerPheIleTyArgLysLeuMetThrPheLeuIleLeuSerIleAlaProValLeu 649
|||
Db 2872 ATATCCCTGATCTACCGCTGGCAATTTGACACTTTTACTCTCAATGTTGTTCCCATCATT 2931
|||
Qy 650 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 669
|||
Db 2932 GCTATAGCAGGAGTGGTGAATGAAATGTTCTCTGGACAAGCGCTGAAGATAGAAG 2991
|||
Qy 670 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 689
|||
Db 2992 GAACCTAGAGGTTCTGGAGAGATCGCTACAGAAGCAATTTGAAACCTTTCCACTGCTGCTC 3051
|||
Qy 690 SerLeuThrArgGluLysAlaPheGluGlnMetTyGluGluMetLeuGlnThrGlnHis 709
|||
Db 3052 TCTTTGACTCGGAGCAGAAAGTTTGAACCTATGTATGCTCCAGAGCTTGCAGATACCATAC 3111
|||
Qy 710 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyArgAlaPheSerHisAlaPhe 729
|||
Db 3112 AGAATGCTTTGAAGAAAGCGCAGCTTTTGGGATCACTTTCTCTTCCCTCCAGGCCATG 3171
|||
Qy 730 IleTyPheAlaTyArgAlaGlyPheArgPheGlyAlaTyLeuIleGlnAlaGlyArg 749
|||
Db 3172 ATGTATTTCTCTATGCTGCTGTTTCCGGTTTGTATGCTGCTTGGTGGCAGCAGAACTC 3231
|||
Qy 750 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyArgLysAlaMetAlaIle 769
|||
Db 3232 ATGACATTTGAAATGTTCTGTTAGTATTTCTGACGTATTTCTTTGTTGGTGGCAGCATG 3291
|||
```

```
Qy 770 GlyLysThrLeuValLeuAlaProGluTySerLysAlaLysSerGlyAlaAlaHisLeu 789
|||
Db 3292 GGGCAGGTCAAGTTCATTCGCTCTCACTACGCGAAAGCAAAAGTCTCGGCATCCACATC 3351
|||
Qy 790 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 809
|||
Db 3352 ATCAGATCATTTGAGAAATCCCTGAGATTGACAGTACAGCAGCGAGGCTTTGAAGCTT 3411
|||
Qy 810 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyProCysArgPro 829
|||
Db 3412 AATATGTTTGAAGGAAATGTAATTTAATGAGTCACTGTTCAACTATCCACCCGACCC 3471
|||
Qy 830 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 849
|||
Db 3472 AACATCCCACTGCTTCAGGGGCTGAGCTTAGAGTGAAGAAAGGCGACGCTGGGCCCTC 3531
|||
Qy 850 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyArgP 869
|||
Db 3532 GTGGCAGCAGTGGCTGCGGGAAGAGTACAGTGTCTCAGCTCTCTTGAGGCTTCTATGAC 3591
|||
Qy 870 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 889
|||
Db 3592 CCCATGGCCGGAACAGTGTCTTAGATGGCAAGAAATAAAGCAACTCAATGTCCAGTGG 3651
|||
Qy 890 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909
|||
Db 3652 CTCGCGCCCACTGGGCATTTGTCCCAAGAGGCCATCTCTGTTGACTGCAGCATCGCC 3711
|||
Qy 910 GluAsnIleAlaTyArgLysAspAsnSerArgValValProLeuAspGluIleLysGluAla 929
|||
Db 3712 GAGAACATTTGCTTACGGACCAACAGCCGTGCTGCTGCTCTATAAGAGATCGTGAAGGCA 3771
|||
Qy 930 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyArgThrGln 949
|||
Db 3772 GCCAAGRGCCCAACATCCACAGTTCATCGACTCACTGCTGAGAAATACAAACACCAGA 3831
|||
Qy 950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 969
|||
Db 3832 GTGGGAGACAAAGGGACTCAGCTGTGCGCGGGCAGAAAGCAGCGCATCGCCATCGCGCGC 3891
|||
Qy 970 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989
|||
Db 3892 GCCTCGTCAGACAGCTCAGATCTTACTTCTGGATGAAGCGACATCAGCTCTGGATACG 3951
|||
Qy 990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009
|||
Db 3952 GAGAGTGAAGAGTGTCTCCAGGAGCGCTGGACAAAGCCAGGGAAGCCGACCTGCATT 4011
|||
Qy 1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1029
|||
Db 4012 GTGATCGCGCAGCGCTGTCCACCATCCAGAACCGCAGACTTGTATCGTGTGTATTCAGAAC 4071
|||
Qy 1030 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyPhe 1049
|||
Db 4072 GGCAGGTCAAGAGCAGCGCACCCACCGAGCTGCTGCCCGAAGAGGCATCTATTTC 4131
|||
Qy 1050 LysLeuValAsnAlaGlnSer 1056
|||
Db 4132 TCGATGCTCAGTGTGCGGCT 4152
|||
RESULT 8
AAZ52048
ID AAZ52048 standard; cDNA; 4425 BP.
XX
AC AAZ52048;
XX
XX 18-JUL-2000 (first entry)
XX
DE Rat multidrug resistance protein 1a cDNA derived from EST sequences.
XX
XX Multidrug resistance protein 1a; mdrla; multi-specific drug transporter;
KW drug formulation; formulation excipient; compound design; inflammation;
KW pharmacokinetic; oral absorption; cancer; cardiovascular disease;
```


KW central nervous system disorder; auto-immune disease; kidney disease;
 KW EST; expressed sequence tag; ss.

OS Rattus rattus.

XX Key Location/Qualifiers
 FH 352..4170
 CDS /*tag= a
 FT /product= "Rat multidrug resistance protein"

XX WO200015650-A1.

PN 23-MAR-2000.

XX 10-SEP-1999; 99WO-US20770.

XX 17-SEP-1998; 98US-0156800.

PR 09-DEC-1998; 98US-0208809.

XX (SMIK) SMITHKLINE BEECHAM CORP.

XX Brun KA, Chenery RJ, Ellens H, Feild JA, Yue L;

XX WPI; 2000-271372/23.

DR P-PSDB; AAY70597.

XX Isolated rat mdrla polynucleotides and polypeptides, useful in assays
 PT to provide information on drug formulation, selection of formulation
 PT excipients and compound design -

XX Claim 11; Page 27-29; 33pp; English.

PS The present cDNA sequence encodes rat multidrug resistance protein 1a
 CC (mdrla). This cDNA is derived from EST (expressed sequence tag)
 CC sequences. Mdrla is used in assays to provide information on
 CC drug formulation, selection of formulation excipients and compound
 CC design. They are used in cell based, membrane based, binding or other
 CC assays to provide information that may enhance drug formulation. This
 CC invention further relates to the generation of in vivo and in vitro
 CC comparison data to predict oral absorption and pharmacokinetics. This
 CC enables the selection of drugs with optimal pharmacokinetics, i.e. good
 CC oral bioavailability, brain penetration, plasma half life, and minimum
 CC drug interaction. Transgenic and knock-out animals created using rat
 CC mdrla provides an insight into treating and preventing human diseases
 CC including cancer, inflammation, cardiovascular disease, central nervous
 CC system disorders, auto-immune and kidney disease.

XX SQ Sequence 4425 BP; 1250 A; 992 C; 1113 G; 1067 T; 2 U; 1 other;

Alignment Scores:

Pred. No.: 6,25e-283 Length: 4425
 Score: 3106.50 Matches: 600
 Percent Similarity: 75.54% Conservative: 206
 Best Local Similarity: 56.23% Mismatches: 218
 Query Match: 58.28% Indels: 43
 DB: 21 Gaps: 7

US-09-873-409-4 (1-1058) x AAZ52048 (1-4425)

QY 4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluVal 23
 Db 1039 TCATTACTGATAGGAACCTCCAGCTTATGCAAAAGCTGGAGCAGTTGCTGAAGAATC 1098
 QY 24 LeuSerSerLeuArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43
 Db 1099 TTAGCAGCCATCAGAACTGTGATTCCTTTGGAGGACAAAAGAGAACTTGA- 1152
 QY 44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63
 Db 1152 -----AspIleLysLysAlaAspGluGlnMetGluSer----- 435
 QY 64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
 Db 1152 -----AspIleLysLysAlaAspGluGlnMetGluSer----- 435

Db 1153 -----AGGTACAAATAACAATTTGGAAAGAGCTAAAGAGGCTGGGATAAAG 1197
 QY 84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
 Db 1198 AAAGCTATCAGGCCAACATTTCCATGGGTGCAGCTTTCTTCTGCTATCTATCATATAT 1257
 QY 104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123
 Db 1258 GCTCTGGCATTTCTGGTATGGGACTTCTTGGTCATCTCAAAAGAA-----TACACTATT 1311
 QY 124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
 Db 1312 GGACAAAGTGTCTCACTGTCTTTTTTCTGTATTAAATGGAGCATTCAGTGTGGCAGGCA 1371
 QY 144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163
 Db 1372 TCTCAAAATATTGAAGCCTTCGCCAATGCTAGAGGAGCAGCTTATGAAGTCTTCAGTATA 1431
 QY 164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
 Db 1432 ATTGATAATAAGCCAGTATAGACAGCTTCTCAAGAGTGGGCACAAACCCGACACATA 1491
 QY 184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
 Db 1492 CRAAGAAATTTGGAAATTCAAAAATATTCACTTCAGTTACCCGCTCGAAAAGACGTTCA 1551
 QY 204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
 Db 1552 ATCTTGAGGGCTCAACCTGAAGGTGAAGCGGACGAGCGGTAGCCCTGGTTGGCAAC 1611
 QY 224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243
 Db 1612 AGTGGCTGTGGGAAAAGCACAACTGTCCAGTGTCTCAGAGGCTCTACGACCCCATAGAG 1671
 QY 244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
 Db 1672 GGCAGGTCTAGTATCGACGACAGGACATCAGGACCAATCAATGTGAGGTATCTGCGGAA 1731
 QY 264 HisIleGlyValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283
 Db 1732 ATCATTTGGGTGTGAGTCAGGAACCGCTGTCTTGGCCACCAATTCGCGAAACATT 1791
 QY 284 LysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303
 Db 1792 CGCTATGGCCGAGAAACGTCACCATGGATGATAGAGAAAGCTGTCAAGGAAGCCAA 1851
 QY 304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyLysGly 323
 Db 1852 GCCTATGATTTCATCATGAAACTGCCCCCAAAATTTGACACCCCTGGTGGTAGAGAGG 1911
 QY 324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343
 Db 1912 GCGCAGCTGAGTGGGGACAGAAACAGAGATCGCAATTCGCCGGCCCTGGTCCGCAAC 1971
 QY 344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysAla 363
 Db 1972 CCCAAGATCCTTTGTTGGATGAGCCACGTCAGCTTGGACACAGAAACGCGCGTG 2031
 QY 364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383
 Db 2032 GTTCAGGCCCTCTGGATAGGCTAGAGAGCGCGACCACTTGTGATAGTACCCGC 2091
 QY 384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403
 Db 2092 TTGTCTACAGTTCGCAATGCTGAUGTCATTTGCTGGTTTGTGATGGTGTGTCATTTGGAG 2151
 QY 404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423
 Db 2152 CAAGGAAATCATGATGAGCTCATGAGAGAGAAAGGAATTTACTTCAAACTTGTCTACT 2211
 QY 424 Gln-----AspIleLysLysAlaAspGluGlnMetGluSer----- 435
 Db 2212 CAGACAGCAGGAAATGAAATTTGAATTGAGGAATGAAGCTTTGTGAATCTAAAGAUGGAATT 2271


```
QY 436 -----MetThrTyrSerThrGluArgLysThrAsnSerLeu-----ProLeuHisSer 451
Db 2272 GATAAATGTGGACATGCTTCAAAAGATTCRGGATCCAGTCTAATAAGAAAGATCAACT 2331
QY 452 ValLysSerIleLysSerAspPheIleAapLysAlaGluLysThr-----GlnSerLys 470
Db 2332 CGCAAAAGCATCGTGGGCCACATGATCAAGACGGGAACCTTAGCACCAAGAGGCTCTG 2391
QY 471 GluIleSerLeuProGluValSerLeuLysIleLeuLysLeuAsnLysPProGluTir 490
Db 2392 GATGACGACGTACTCCAGCTTCTTTTGGCGGATCTCTGAAGTTGAATTCAACTGAATGG 2451
QY 491 ProPheValValLeuGluThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 510
Db 2452 CCTATTATTTGTGGTGTGATTTGTGTCATATAAATGAAGGCTTGCACACGACATTC 2511
QY 511 SerIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThrLeu--- 529
Db 2512 TCCATAATATTTTCAAGGTTGTAGGGTTTTTACAAAATAATGACACCCCTGAAATCCAG 2571
QY 530 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 549
Db 2572 CGCACAACAGCAACTGTGTTCTTATTGTTCTGATCTCTGGATCATCTCTTTCATT 2631
QY 550 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 569
Db 2632 ACCTTTTCTCTCAGGCTTTCACATTTGGCAAGCTGGAGAGATCTCTCAACAGGCACTC 2691
QY 570 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTirPheAspGluLysGlu 589
Db 2692 CGATACATGCTTCAAAATCCATGCTGAGACAGGACATAGCTGGTTGATGACCCCTAAA 2751
QY 590 AsnSerThrGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 609
Db 2752 AACCCACAGGCGCTGACCAACAGCTTGGCAATGACCTCTCAAGTGAAGGGGCT 2811
QY 610 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629
Db 2812 ACAGGCTCTAGGCTTGCTGTTATTACCAGACACATAGCAAAATCTTGGGACAGCATCATC 2871
QY 630 IleSerPheIleTyrGlyTirPleuMetThrPheIleLeuSerIleAlaProValLeu 649
Db 2872 ATATCCCTGATCTACGGCTGGCAATGGACACTTTTACTCTCAGCAATTTGTTCCTCAT 2931
QY 650 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 669
Db 2932 GCTATAGCAGGAGTGTGTAATGAAATGTTGTCTGGCAAGCGCTGAAGATAGAAG 2991
QY 670 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 689
Db 2992 GAATAGAGGTTCTGGAGAGATCGTACAGAGCAATTCGAAACTTTCGCACTGTGCTC 3051
QY 690 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluMetLeuGlnThrGlnHis 709
Db 3052 TCTTTGACTCGGAGCAGAAAGTTTGAACATATGATGCCAGAGCTTGCAGATACCATAC 3111
QY 710 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 729
Db 3112 AGAAATGCTTTGAAGAAAGCGACGCTTTGGGATCATCTTTCTCTTCAACCCAGGCGATG 3171
QY 730 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrIleLeuIleGlnAlaGlyArg 749
Db 3172 ATGTAATTCCTATGCTGCTGTTGTTCCGGTTTGATGCTTCTTGGGACGAGAACTC 3231
QY 750 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 769
Db 3232 ATGACATTTGAAATGTTCTGTAGTATTCTCAGTATTGCTTGTGGTGCATGGCAGTG 3291
QY 770 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 789
Db 3292 GGGCAGGTCAGTTTCATTCGCTCTCTGACTACGCGAAGAAAGCCAAAGTCTCGGCATCCCATC 3351
```

```
QY 790 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 809
Db 3352 ATCAGGATCATTTGAGAAAAATCCCTGAGATTGACAGCTACAGCAGCGGAGGCTTTGAAGCCT 3411
QY 810 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 829
Db 3412 AATATGTTGAGAGAAATGTGAAATTTAATGAGTTCATGTTCAACTATCCACCCGACCC 3471
QY 830 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 849
Db 3472 AACATCCAGTGTCTCAGGGCTGAGCTTAGAGGTGAAGAAAGGCGAGCGTGGCCCTC 3531
QY 850 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 869
Db 3532 GTGGGACAGAGTGGCTGCGGAGAGTACAGTGGTCCAGCTGCTTGAGCGCTTCTATGAC 3591
QY 870 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTir 889
Db 3592 CCATGGCCGGAACAGTGTCTTAGATGGCAAGAAATAAGCAACTCAATGTCCAGTGG 3651
QY 890 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909
Db 3652 CTCGGGCCACCTGGGCTTGTGCCAGGAGCCCATCTCTGTTGACTGCAGCATGCC 3711
QY 910 GluAsnIleAlaTyrGlyAspAsnSerArgValProLeuAspGluIleLysGluAla 929
Db 3712 GAGAACATTCCTTACGAGACACAGCCGCTGCTCTCATAGGAGATCGTGAAGGCA 3771
QY 930 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 949
Db 3772 GCCAAGGAGCCCAACATCCACAGTTCATCGACTCATCTGAGAAATACAAACACAGA 3831
QY 950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 969
Db 3832 GTGGGACAGAAAGGACTCAGCTGTGGGGGCGGACAGCGCATCGCCATCGCGCG 3891
QY 970 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989
Db 3892 GCCTCGTCAGACAGCTCACATCTTACTTCTGGATGAAGCGACATCAGCTCTGGATACG 3951
QY 990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009
Db 3952 GAGAGTGAAGAGTCTCGAGGAAGCGCTGGCAAGCGGAGGAGGAGGAGGAGGAGGAGGAG 4011
QY 1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1029
Db 4012 GTGATCGCGCACCGCTCTCCACCATCCAGACGCGAGCTTGATCGTGGTGTTCAGAAC 4071
QY 1030 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1049
Db 4072 GGCAGGTCAAGGAGCACGGCACCCACAGAGCTGTGGCCCAAGAAAGGATCTATTTC 4131
QY 1050 LysLeuValAsnAlaGlnSer 1056
Db 4132 TCGATGCTCAGTGTGAGGCT 4152
RESULT 9
AAZ49332
ID AAZ49332 standard; cDNA; 3860 BP.
XX
AC AAZ49332;
XX
DT
XX
XX 14-MAR-2000 (first entry)
XX Human wild-type multidrug resistance-1 (MDR-1) cDNA.
XX Multidrug resistance; MDR-1; P-glycoprotein;
XX transmembrane efflux pump; haematopoietic stem cell; transduction;
XX bone marrow transplantation; chemotherapy; radiation therapy; cancer;
XX gene therapy; gene replacement; genetic defect; thalassemia;
XX Gaucher's disease; sickle cell anaemia; leukaemia; ex vivo expansion;
XX cytokine; wild-type; ds.
```


OS Homo sapiens.
 XX Location/Qualifiers
 FH 1..3843
 FT /tag= a
 FT /product= "Human wild-type MDR-1 protein"
 FT replace (553..555, GTT)
 FT /tag= b
 FT mutation
 FT /note= "cDNA sequence of G185V human mutant MDR-1 given
 XX in AAZ49333"
 PN WO9961589-A2.
 XX
 PD 02-DEC-1999.
 XX
 XX 27-MAY-1999; 99WO-US11825.
 XX
 PR 28-MAY-1998; 98US-0086988.
 XX
 XX (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
 XX
 XX Sorrentino B, Bunting K;
 XX
 DR WPI; 2000-072615/06.
 DR P-PSDB; AAY58186.
 XX
 PT Ex vivo expansion of hematopoietic stem cells transduced with a
 PT sequence encoding human multidrug resistance-1, used for bone marrow
 PT transplantation -
 XX
 PS Claim 10; Page 68-70; 113pp; English.
 XX
 CC This sequence represents cDNA encoding human wild-type
 CC multidrug resistance protein MDR-1. MDR-1 is a transmembrane
 CC efflux pump, responsible for the export of drugs from cells,
 CC particularly cancer cells. Wild-type MDR-1 shows increased
 CC resistance to etoposide and decreased resistance to vinca
 CC alkaloids compared with a mutant form (AAY58187) where the Gly at
 CC position 185 is replaced by Val. The invention relates to transducing
 CC hematopoietic stem cells with nucleic acid encoding an MDR protein
 CC and culturing the modified cells. The modified hematopoietic stem
 CC cells are useful in bone marrow transplantation (to reconstitute
 CC hematopoietic systems in patients who have undergone chemotherapy or
 CC radiation therapy) and in ex vivo gene therapy of genetic defects in
 CC cells derived from hematopoietic stem cells, e.g., thalassemia,
 CC Gaucher's disease, sickle cell anemia or leukaemia. The modified
 CC cells can also be used to identify factors involved in regulating
 CC proliferation and differentiation in hematopoietic stem cells.
 CC Hematopoietic stem cells that express MDR-1 will be protected against
 CC chemotherapeutic agents, so can be engrafted while the patient is
 CC undergoing chemotherapy. Expansion of (rare) hematopoietic stem cells
 CC provides sufficient cells to permit standard biochemical analysis.
 CC Overexpression of MDR-1 allows cytokine-driven expansion of
 CC hematopoietic stem cells by at least 10-fold compared with a maximum
 CC of 4-fold in known procedures.
 XX
 SQ Sequence 3860 BP; 1135 A; 746 C; 957 G; 1022 T; 0 other;
 XX

Alignment Scores:
 Pred. No.: 1.88e-282 Length: 3860
 Score: 3100.50 Matches: 595
 Percent Similarity: 75.54% Conservative: 211
 Best Local Similarity: 55.76% Mismatches: 218
 Query Match: 58.17% Indels: 43
 DB: 21 Gaps: 6

US-09-873-409-4 (1-1058) x AAZ49332 (1-3860)

QY 4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluVal 23
 Db 712 TCATTTACTGATTAAGAACTCTTAGCGTATGCAAAAGCTGGAGCAGTAGCTGAAGAGTC 771
 QY 24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43

Db 772 TTGGCAGCAATTAGAACTGTGATTGCTATTGGAGGACAAAGAAAGAACTTGAA----- 825
 QY 44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63
 Db 825 ----- 825
 QY 64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
 Db 826 -----AGGTACACAAATAATTAGAGAGAGCTAAAGAAATTGGGATAAG 870
 QY 84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
 Db 871 AAAGCTATTACAGCCAATATTCTATAGGTGCTGCTTCTCTGCTGATCTATGATCTTAT 930
 QY 104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123
 Db 931 GCTCTGGCCTTCGTGATGGGACCACTTGGTCTCTTCAGGGGAA-----TATTCTATT 984
 QY 124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
 Db 985 GGACAAGTACTCACTGTTCTTTCTGTATTAAATTGGGGCTTTTAGTGTGGACAGGCA 1044
 QY 144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163
 Db 1045 TCTCCAAGCATTTGAAGCATTTGCAAAATGCAAGAGAGCAGCTTATGAAATCTTCAAGATA 1104
 QY 164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
 Db 1105 ATTGTAATAAGCCAAGTATTGACAGTATTCGAAGAGTGGGCACAAACAGATAATATT 1164
 QY 184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
 Db 1165 AAGGGAATTTGGAATTCAGAAATGTTCACTTCAGTTACCATCTCGAAAGAAGTTAAG 1224
 QY 204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
 Db 1225 ATCTGAAGGGCTGAACCTGAAGGTGCAGAGTGGCAGACGGTGGCCCTGGTTGGAAC 1284
 QY 224 AsnGlySerGlyLysSerThrValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243
 Db 1285 AGTGGCTGTGGGAAGAGCACACAGTCCAGCTGATGCAGAGGCTCTATGATCCCCACAG 1344
 QY 244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
 Db 1345 GGGATGGTCAGTGTTCATGGACAGATATTAGACCATAATGTAAGGTTTCTACGGAA 1404
 QY 264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283
 Db 1405 ATCATTTGGTGTGAGTCAGGAACCTGTATTGTTGCCACCACGATAGCTGAAAACATT 1464
 QY 284 LysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303
 Db 1465 CGTATGGCCGTGAAATGTCACCATGGATGAGATGAGAAAGCTGTCAAGGAGGCAAT 1524
 QY 304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323
 Db 1525 GCCTATGACTTTATCATGAACTGCCTCATAAATTTGACACCTCGTTGGAGAGAGGGG 1584
 QY 324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343
 Db 1585 GCCCAGTTGAGTGGTGCGCAGAAAGCAGAGGATCGCCATTCAGCTGCGCTTGGCAAC 1644
 QY 344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
 Db 1645 CCCAAGATCCTCTGCTGGATGAGGCCACGTCAGCTTGGACACAGAAACGGAAGCAGTG 1704
 QY 364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383
 Db 1705 GTTCAGGTGGCTCTGGATTAAGGCCAGAAAGGTGGACCAACCATTTGTAGTCTCATCGT 1764
 QY 384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403

Db 1765 TTGCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTCGATGATGAGTCATTTGGAG 1824
Qy 404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423
Db 1825 AAGGAAATCATGATGAATCATGAAAGAAAGGCAATTTACTTCAAACTGTGCACAATG 1884
Qy 424 Gln-----AspIleLysLysAlaAspGlnMetGluSerMetThr 437
Db 1885 CAGACAGCAGGAAATGAAGTTGAATTAGAAATGACAGTCGATGAATCCAAAGTGAAT 1944
Qy 438 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal----- 452
Db 1945 GATGCTTGGAAATGCTTCAAAATGATTCAGATCCAGTCTTAATAGAAAGATCAACT 2004
Qy 453 ---LysSerIleLys---SerAspPheIleAspLysAlaGluGlnSerThrGlnSerLys 470
Db 2005 CGTAGGAGTGCCTGGATCACAAGCCCAAGACAGAAAGCTTAGTACCAAGAGGCTCTG 2064
Qy 471 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuLeuLysProGluTyr 490
Db 2065 GATGAAAGTATACCTCCAGTTCTCTTTTGAGGATATGAAGCTAAATTTAACTGAATGG 2124
Qy 491 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 510
Db 2125 CCTATTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2184
Qy 511 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 529
Db 2185 GCATAATATTTCAAGATATATAGGGTTTTTACAGAATGATGATCTCGAAACAAAA 2244
Qy 530 LysHisAspAlaGluLeuTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 549
Db 2245 CGACAGAAATAGTAATCTGTTTCTCACTATTGTTCTAGCCCTTGGAAATATTTCTTTATT 2304
Qy 550 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyLeuLeuLeuThrMetArgLeu 569
Db 2305 ACATTTTCTTCAAGGTTTTCACATTTGGCAAAAGCTGGAGAGATCTCCACCAAGCGCTC 2364
Qy 570 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGluLysGlu 589
Db 2365 CGATACATGTTTCCGATCCGCTCAGACAGGATGTGAGTTGGTTGATGACCTTAA 2424
Qy 590 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 609
Db 2425 AACACCACTCGAGCATTGCTACAGGCTCGCCAATGATGCTCAAGTTAAAGGGCT 2484
Qy 610 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629
Db 2485 ATAGGTTCCAGGCTTGCTGTAATTTACCCAGAAATATAGCAAAATCTTGGGACAGGAATATT 2544
Qy 630 IleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeu 649
Db 2545 ATATCTTTCATATGTTGGCAACTAACACTGTTTACTCTTAGCAATGTGACCCATCAT 2604
Qy 650 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 669
Db 2605 GCAATACAGGAGTGTGTGAATGAAATGTTGCTCGCAAGCACTGAAAGATAAGAAA 2664
Qy 670 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 689
Db 2665 GAAC TAGAAGGTGCTGGGAAAGATCGCTACTGAAAGCAATAGAAAACTTCCGAACCGTTGTT 2724
Qy 690 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 709
Db 2725 TCATTGACTCAGGAGCAGAAAGTTTGAACATATGATGCTCAGAGTTTGCAGGTTACCATAC 2784
Qy 710 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 729
Db 2785 AGAACTCTTTGAGAAAGCACACATCTTTGGAATTACATTTTCTTCCACCAGCAATG 2844
Qy 730 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 749
Db 2845 ATGTATTTTCTTCTATGCTGGATGTTTCCGGTTTGGAGCTTACTTGGTGCCACATAAACTC 2904

Qy 750 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 769
Db 2905 ATGAGCTTTGAGGATGTTCTGTAGTATTTCAGCTGTTGTCTTTGGTGCATGGCCGTG 2964
Qy 770 GlyLysThrLeuValLeuAlaProGluTyrSerLysLysSerGlyAlaAlaHisLeu 789
Db 2965 GGGCAAGTCAGTTCAATTTGCTCTGCTATGTCACAAAGCCAAATATATCAGCAGCCACATC 3024
Qy 790 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 809
Db 3025 ATCATGATCATTTGAAAAAACCCCTTTGATTGACAGTACAGCAGGAGGCCTTAATGCCG 3084
Qy 810 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPheTyrProCysArgPro 829
Db 3085 AACACATTGGAAGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCACCCGACCG 3144
Qy 830 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 849
Db 3145 GACATCCCACTGCTTCAGGACTGAGCCTGGAGGTGAAGAAAGGCGCAGACGCTGGCTCTG 3204
Qy 850 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 869
Db 3205 GTGGGACAGCAGTGGCTGTGGGAAGACACAGTGTCTCCAGCTCTCTGGAGCGGTTCTACGAC 3264
Qy 870 ProValGlnGlyGlnValLeuPheAspGlyValAlaAspAlaLysGluLeuAsnValGlnTyr 889
Db 3265 CCCTTGGCAGGAAAGTGTCTGTGATGTCAAAGAAATAAAGCGACTGAATGTTCACTGG 3324
Qy 890 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909
Db 3325 CTCGAGACACCTGGGCATCGTGTCCAGGAGCCCATCTCTGTTTCACTGCACGACATGCT 3384
Qy 910 GluAsnIleAlaTyrGlyAspAsnSerArgValProLeuAspGluIleLysGluAla 929
Db 3385 GAGAACATTGCCATTGAGACAAACACCGCGGTGTGTCTCACAGGAGAGATCTGTGAGGGCA 3444
Qy 930 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 949
Db 3445 GCAAGGAGGCCCAACATACATGCTCTTCATCGAGTCACTGCCTAATAATAATATAGCACTAAA 3504
Qy 950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 969
Db 3505 GTAGGAGACAAAGGAACTCAGCTCTCTGTGGTGGCCAGAAAACAACGCAATGCTAGTCTGT 3564
Qy 970 AlaLeuLeuGlnLysProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989
Db 3565 GCCCTTGTAGACAGCTCATATTTTGTGATGAAGCCACGTCAGCTCTGGATACA 3624
Qy 990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009
Db 3625 GAAAGTGAAAGGTTGTCCAAGAAAGCCCTGGACAAAGCCAGAGAGCGCCGACCTGCATT 3684
Qy 1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1029
Db 3685 GTGATTGCTCACCGCTGTCCACCATCCAGATGTCAGACTTAATAGTGGTGTTCAGAAT 3744
Qy 1030 GlyLysIleLysGluGlnGlyThrHisGlnLeuLeuArgAsnArgAspIleTyrPhe 1049
Db 3745 GGCAGAGTCAAGGAGCATGGCAGCATCAGCAGCTGCTGGCACAGAAAGCACTATTTT 3804
Qy 1050 LysLeuValAsnAlaGlnSer 1056
Db 3805 TCAATGTCAGTGTCCAGGCT 3825
RESULT 10
AAZ49333
ID AAZ49333 standard; cDNA; 3860 BP.
XX
AC AAZ49333;
XX
DT 14-MAR-2000 (first entry)
XX

Human G185V mutant multidrug resistance-1 (MDR-1) cDNA.

Multi:drug resistance; MDR-1; P-glycoprotein;
transmembrane efflux pump; haematopoietic stem cell; transduction;
bone marrow transplantation; chemotherapy; radiation therapy; cancer;
gene therapy; gene replacement; genetic defect; thalassaemia;
Gaucher's disease; sickle cell anaemia; leukaemia; ex vivo expansion;
cytokine; mutant; ds.

Synthetic.

OS Homo sapiens.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Key Location/Qualifiers
CD5 1.3843
/*tag= a
/product= "Human G185V mutant MDR-1 protein"

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Alignment Scores:
Pred. No.: 1.88e-282 Length: 3860
Score: 3100.50 Matches: 595
Percent Similarity: 75.54% Conservative: 211
Best Local Similarity: 55.76% Mismatches: 218
Query Match: 58.17% Indels: 43
DB: 21 Gaps: 6

US-09-873-409-4 (1-1058) x AAZ49333 (1-3860)

Sequence 3860 BP; 1135 A; 746 C; 956 G; 1023 T; 0 other;

QY 4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValaGluGluVal 23
Db 712 TCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCTGGAGCAGTAGCTGAAGAGGTC 771
QY 24 LeuSerSerIleAArgThrValIleAlaPheAArgAlaGlnGlnLysGluLeuGlnAArgSer 43
Db 772 TTGGCAGCAATTAGAAGCTGCTGATTTGCAATTTGGAGGCAAAAGAAAGAACTTGAA----- 825
QY 44 PheLeuLeuAsnIleThrArgTyrAlaTyrPheTyrPheProGlnTrpLeuLeuSerCys 63
Db 825 ----- 825
QY 64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
Db 826 -----AGGTACACAAATAATTTAGAAGAAGCTAAAGAAATTTGGATAAAG 870
QY 84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
Db 871 AAAGCTATTACAGCAATATTCTTATAGGTGCTGCTTCTGCTGATCTATGCTATCTAT 930
QY 104 GlyLeuAlaPheThrTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123
Db 931 GCTCTGGCTTCTGGTATGGACCACTTGGTCTCTCAGGGGAA-----TATTCTATT 984
QY 124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
Db 985 GGACAAGTACTCACTGATTTCTTCTGATTAAATTTGGGGCTTTTAGTGTGGACAGGCA 1044
QY 144 ValProHisPheGluThrPheAlaIleAlaAArgGlyAlaAlaPheHisIlePheGlnVal 163
Db 1045 TCTCCAAGCATTTGAAGCATTTGCAATTTGCAAGAGGAGCAGCTTATGAAATCTTCAAGTA 1104
QY 164 IleAspLysIleProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
Db 1105 ATTGATAAATGAAGCAAGTATTGACAGCTATTGCAAGAGTGGGCACAAACAGATAATTT 1164
QY 184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
Db 1165 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTACCCTCTCGAAAGAAAGTTAAG 1224
QY 204 IleLeuLysGlyLeuAsnLeuAArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
Db 1225 ATCTTGAAGGGCTGAACCTGAAGTGCAGTGGCAGCAGCGTGGCCCTGGTGGAAAC 1284
QY 224 AsnGlySerGlyLysSerThrValValGlnLeuGlnAArgLeuTyrAspProAspAsp 243
Db 1285 AGTGGCTGTGGGAAGAGCACACAGTCCAGCTGATGAGAGGCTCTATGACCCACAGAG 1344
QY 244 GlyPheIleMetValAspGluAsnAspIleAArgAlaLeuAsnValArgHisTyrArgAsp 263
Db 1345 GGGATGGTCAGTCTTGTATGGACAGGATATTAGCCATTAATGTAAGGTTTCTACGGAA 1404
QY 264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283
Db 1405 ATCATTTGGTGTGGTGTAGTCAAGAACCTGTATTGTTGCCACCACGATAGCTGAAACATT 1464
QY 284 LysTyrGlyArgAspValThrAspGluGluMetGluAArgAlaAlaAArgGluAlaAsn 303
Db 1465 CGGTATGGCGGTGAAATGTCCACCATGTGATGAGATGAGAAAGCTGTCAAGGAAGCCAA 1524
QY 304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323
Db 1525 GCCTATGACTTTTATCATGAAACTGCCTCATATAATTTGACACCCCTGGTTGGAGAGAGGG 1584
QY 324 AlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343
Db 1585 GCCCAGTTGATGTGGGCAGAAAGCAGAGGATGCCATTCAGCTGCGCTGGTTCGCAAC 1644
QY 344 ProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
Db 1645 CCCAAGATCTCTCTGTGGATGAGGCCACCTGAGCTTTGGACACAGAAAGCGAAGCAGTG 1704

QY 364 ValGlnAlaLeuGluLysAlaSerLysGlyArgThrThrIleValAlaHisArg 383
Db 1705 GTTCAGGTGCTCTGGATAGGCGAGAAAGGTCGGACCACTTGTGATAGCTATCGT 1764
QY 384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403
Db 1765 TTGTCTACAGTTCTGTAATGCTGAGCTCATCGCTGTTTCGATGATGGAGTCATGTGGAG 1824
QY 404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423
Db 1825 AAAGGAATCATGATGAACATCATGAAGAGAAAGGCACTTACTTCAAACCTGTGCACAAATG 1884
QY 424 Gln-----AspIleLysAlaAspGluGlnMetGluSerMetThr 437
Db 1885 CACAGCAGCAAGAAATGAAGTTGAATTAGAAATGACAGCTGATGAATCCAAAGAGTGAAT 1944
QY 438 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal----- 452
Db 1945 GATGCCTTGGAAATGCTTCAATGATTCAAGATCCAGTCTAATAGAAAGAAATCAACT 2004
QY 453 ---LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys 470
Db 2005 CGTAGGAGTCTCGGTGATCACAAGCCCAAGACAGAAAGCTTAGTACCAAGAGGCTCTG 2064
QY 471 GluIleSerLeuProGluValSerLeuLysIleLeuLysLeuAsnLysProGluTyr 490
Db 2065 GATGAAGATATACCTCCAGTCTTCTTTGGAGGATTATGAAGCTAAATTTAACTGAATGG 2124
QY 491 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 510
Db 2125 CCTATTTTGTGTGTGTATTTTGGCCATTATTAATGGAGCCCTGCAACAGCAATTT 2184
QY 511 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 529
Db 2185 GCAATATATTTCAAAGATTATAGGGTTTTACAGAAATTTGATGATCCTGAAACAAA 2244
QY 530 LysHisAspAlaGluIleLysThrMetIlePheValIleLeuGlyValIleCysPheVal 549
Db 2245 CGACAGAAATAGTAATCTGTTTCACTATTTCTTACCCCTTGGAAATTTATTTCTTTAT 2304
QY 550 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 569
Db 2305 ACATTTTCTTCAAGTTTCACTTTTGGCAAGCTGGAGAGTCTTCAACCAAGCGGCTC 2364
QY 570 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGlyLysGlu 589
Db 2365 CGATACATGTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTGATGACCTAAA 2424
QY 590 AsnSerThrGlyGlyLeuThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 609
Db 2425 AACACCACTGGAGCATTTGACTACAGGCTCGCAATGATGCTCAAGTTAAAGGGCT 2484
QY 610 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629
Db 2485 ATAGTTTCCAGGCTTGTGTAATTTACCAGAAATATAGCAATCTTGGGACAGAAATATT 2544
QY 630 IleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeu 649
Db 2545 ATATCTTCACTATGTTGGCAACTTAACACTGTACTCTTAGCAATGTACCAATCAT 2604
QY 650 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 669
Db 2605 GCAATAGCAGGAGTTGTTGAATGMAATGTTGCTGGCAAGCACTGAAAGATNAGAAA 2664
QY 670 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 689
Db 2665 GAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAACTTCCGAACCGTGT 2724
QY 690 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuThrGlnHis 709
Db 2725 TCTTTGACTCAGAGCAGAGATTGGAACATATGATGCTCAGAGTTTGCAGGTACCAATAC 2784
QY 710 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 729

Db 2785 AGAAATCTTTTGGAAAGCACACATCTTTTGGAAATTTACATTTTCTTCAACCCAGGCAATG 2844
QY 730 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyValaTyrLeuIleGlnAlaGlyArg 749
Db 2845 ATGATATTTTCTATGCTGATGCTTTCCGGTTTGGAGCCTACTTGGTGGCACAATAAATC 2904
QY 750 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 769
Db 2905 ATGAGCTTTGAGGATGTTCTGTGTAGTATTTTCACTGTGTGTCTTGTGTCATGGCCGTG 2964
QY 770 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 789
Db 2965 GGCACAGTCAGTTTCTGCTGACTATGTCACAAAGCCAAATATATCAGCAGCCACATC 3024
QY 790 PheAlaLeuGluLysLysProAsnIleAspSerArgSerGlnGlyLysLysPro 809
Db 3025 ATCATGATCATTTGAAAAACCCCTTTGATGTAGCAGTACAGCAGGAGCCCTAATGCCG 3084
QY 810 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 829
Db 3085 AACACATTGGAAGAAATGTACATTTGGTGAAGTTGTATTCAACTATCCACCCGACCG 3144
QY 830 AspValPheIleLeuArgGlyLeuSerIleGluArgGlyLysThrValAlaPhe 849
Db 3145 GACATCCCATGCTTCCAGGACTGAGCTGGAGGTGAAGAGGCCAGCGCTGGCTCTG 3204
QY 850 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuGluGlnArgLeuTyrAsp 869
Db 3205 GTGGGACAGAGTGGCTGTGGGAAGACACAGTGTCTGCTGAGCGGTCTTACGAC 3264
QY 870 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTyr 889
Db 3265 CCCTTGGCAGGAAAGTGTCTGTATGTGCAAAAGATAAAGCGACTGAATGTTCAGTGG 3324
QY 890 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909
Db 3325 CTCGAGCACACCTGGGCATCGTGTCCAGGAGCCCATCTCTGTTGACTGCACATGCT 3384
QY 910 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 929
Db 3385 GAGAAATTCCTATGAGACACACAGCCGGTGTGTACAGGAAGAGATCTGTGAGGCA 3444
QY 930 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 949
Db 3445 GCAAGGAGGCCCAACATACATGCTTTCATCGACTCTGCTTAATAATATAGCACTAAA 3504
QY 950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlnLysGlnArgLeuAlaIleAlaArg 969
Db 3505 GTAGGAGACAAAGAACTCAGCTCTCTGTGTGGCCAGAAACACGCAATGTCATAGCTCG 3564
QY 970 AlaLeuLeuGlnLysProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989
Db 3565 GCCCTTGTAGACGCTCATATTTTGTGTAAGAACCCACGCTCAGCTCTGGATACA 3624
QY 990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009
Db 3625 GAAAGTGAAAGGTTGTCACAGAGCCCTGGACAAAGCCAGAGAGCCGACCTGCATT 3684
QY 1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1029
Db 3685 GTGATTGCTCACCGCTGTCCACCATCCAGAAATGCAGACTTAATAGTGTGTTCAGAAT 3744
QY 1030 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuArgAsnArgAspIleTyrPhe 1049
Db 3745 GGCAGAGTCAAGAGCATGGCAGCATCAGCTGCTGGCACAGAAAGCATCTATTTT 3804
QY 1050 LysLeuValAsnAlaGlnSer 1056
Db 3805 TCATGCTGATGCTCCAGCT 3825
RESULT 11
ABA94365

ID ABA94365 standard; DNA; 3860 BP.

```

XX AC ABA94365;
XX DT 26-MAR-2002 (first entry)
XX DE Human BCRP DNA related seq id No. 1.
XX DE 31-MAY-2000; 2000US-0584586.
XX DE 29-MAY-2001; 2001US-0866866.
XX DE (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX DE Sorrentino B, Schuetz J;
XX DE WPI; 2002-114368/15.
XX DE P-PSDB; ABB07266.
XX DE Identifying a stem cell, for treating e.g., muscular dystrophy,
XX DE myocardial infarction, Parkinson's disease, or neurodegenerative
XX DE disorders, comprises detecting the expression of an ATP transport
XX DE protein (BCRP) by a cell -
XX DE Disclosure; Page 53-55; 87pp; English.
XX DE
XX DE The invention provides a method of identifying and/or isolating a stem
XX DE cell that involves detecting the expression of an ATP transport protein
XX DE containing a conserved ATP-binding cassette (BCRP) by a cell in a sample
XX DE comprising stem cells. The isolated stem cells may be used in the
XX DE treatment of diseases such as muscular dystrophy, degenerative liver
XX DE disorder, myocardial infarction, Parkinson's disease, degenerative
XX DE disorders of the brain, and for tissue regeneration or replacement.
XX DE Haematopoietic cells can be used in bone marrow transplants (e.g., for
XX DE treatment of leukemia) and for ex vivo gene therapy for treating blood
XX DE diseases such as sickle cell anemia and thalassemia. The stem cells can
XX DE also be used as cell targets in gene therapy protocols. The present
XX DE sequence represents a sequence related to the BCRP for which no relevant
XX DE information has been provided in the specification.
XX DE
XX DE Sequence 3860 BP; 1135 A; 746 C; 957 G; 1022 T; 0 other;
XX DE
XX DE Alignment Scores:
XX DE Pred. No.: 1.88e-282 Length: 3860
XX DE Score: 3100.50 Matches: 595
XX DE Percent Similarity: 75.54% Conservative: 211
XX DE Best Local Similarity: 55.76% Mismatches: 218
XX DE Query Match: 58.17% Indels: 43
XX DE DB: 24 Gaps: 6
XX DE
XX DE US-09-873-409-4 (1-1058) x ABA94365 (1-3860)
XX DE
XX DE 4 SerLeuThrSerLysGluLeuSerAlaThrSerLysAlaGlyAlaValAlaGluVal 23
XX DE 712 TCATTACTGATAAAGAACTCTTAGCGTATGCCAAGCTGAGCAGTAGCTGAAGAGTC 771
XX DE 24 LeuSerSerLysAlaThrValAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43
XX DE

```

```

Db 772 TTGGCAGCAATTAGAACTGTGATTCGATTTGGAGGACAAAGAAAGAACTTGA----- 825
QY 44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63
Db 825 ----- 825
QY 64 ValLeu**PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
Db 826 -----AGGTACAAACAAATTTAGAAAGAGCTAAAGAAATTGGGATAAAG 870
QY 84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
Db 871 AAAGCTATTACAGCAATATTTCTATAGTGTCTCTTCTCTCTCTCTCTCTCTCTCTAT 930
QY 104 GlyLeuAlaPheThrTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123
Db 931 GCTCTGGCCTTCGGTATGGACCACTTGGTCTCTCTCAGGGAA-----TATTCTATT 984
QY 124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
Db 985 GGACAAAGTACTCTACTCTTCTCTTCTCTTCTCTTCTCTTCTCTCTCTCTCTCTCTCT 1044
QY 144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163
Db 1045 TCTCCAAGCATTGAAGCATTTCGAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1104
QY 164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
Db 1105 ATTCATANTAAAGCAAGTATTGACAGCTATTGCAAGAGTGGGACCAACCCAGATAATATT 1164
QY 184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
Db 1165 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCCAGTATCCCATCTCGAAAGAAAGTTAAG 1224
QY 204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
Db 1225 ATCTTGAAGGGCTGAACCTGAAGTGAAGTGGGACAGCGTGGCCCTGGTGGAAAC 1284
QY 224 AsnGlySerGlyLysSerThrValValGlnLeuGlnArgLeuTyrAspProAspAsp 243
Db 1285 AGTGGCTGTGGGAAGAGCACACAGTCCAGCTGATGAGAGGCTCTATGACCCACAGAG 1344
QY 244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
Db 1345 GGGATGTCAGTCTTCATGGACAGGATATTAGGACCAATAATGTAAGTTTCTACGGGAA 1404
QY 264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283
Db 1405 ATCATTTGGTGTGGTGAAGTCAAGAACTGTATTGTTGCCACCAAGTAGCTGAAACATT 1464
QY 284 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303
Db 1465 CGCTATGGCGTGAATATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAT 1524
QY 304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323
Db 1525 GCCTATGACTTTATCATGAACTGCCTCATAAATTTGACACCCCTGGTTGGAGAGAGAGG 1584
QY 324 AlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343
Db 1585 GCCCAGTTGAGTGTGGCAGAAAGCAGAGGATCGCCATTCGACGTGCGCTTGGCAAC 1644
QY 344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
Db 1645 CCCAAGATCTCTCTGATGAGGCCACGTCAGCTTGGACACAGAAAGCCAGCAAGCGAGT 1704
QY 364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383
Db 1705 GTTCAGTGTGCTCTGGATAAGGCCAGAAAGGTCGGACCAACCATTTGTAGTCTCATCGT 1764
QY 384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403
Db 1765 TTGCTACAGTTTCTGTAATGCTGACGTCATCGCTGTTTTCGATGATGAGTATTGTGGAG 1824

```



```
QY 404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423
D 1825 AAGAGAAATCATGATGAATCATGAAAGAAAGAGGCAATTTACTTCAAACTTGTCACAAATG 1884
QY 424 Gln-----AspIleLysLysAlaAspGluGlnMetGluSerMetThr 437
D 1885 GACACAGCAGGAAATGAAGTTGAATGTAAGAAATGACAGCTGATGAATCCAAAGTGAATTT 1944
QY 438 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal----- 452
D 1945 GATGCTTGGAAATGCTTCAATATGATTCAAGATCCAGTCTAATAGAAAGAAAGATCAACT 2004
QY 453 ---LysSerIleLys---SerAspPheIleAspLysAlaGluGlnSerThrGlnSerLys 470
D 2005 CGTAGGAGTGTCCGTGGATCACAGGCCAACAGACAGAAAGCTTAGTACCACCAAGAGGCTCTG 2064
QY 471 GluIleSerLeuProGluValSerLeuLysIleLeuLysLeuAsnLysProGluTyr 490
D 2065 GATGAAAGTATACCTCCAGTTTCTTTGGAGGATTATGAAGCTAAATTTAACTGTAATGG 2124
QY 491 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 510
D 2125 CCTATTGTTGTTGGTGTATTTTGTGTCATTTATTAATGAGGCTGCAACAGCAATTT 2184
QY 511 SerIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 529
D 2185 GCAATAATATTTCAAGATTATAGGGTTTTTACAGAAATTTGATGATCCTGGAACACAAA 2244
QY 530 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 549
D 2245 CGACAGATAGTAACATGTTTTCATATTGTTTCTAGCCCTTGAATTTATTTCTTTTATT 2304
QY 550 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 569
D 2305 ACATTTTCTTCAAGGTTTCATATTGGCAAGCTGGAGAGATCTCCACCAAGCCGGCTC 2364
QY 570 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaThrPheAspGluLysGlu 589
D 2365 CGATACATGGTTTTCGATCCATGCTCAGACAGGATGTGAGTTGGTTGATGACCCTAAA 2424
QY 590 AsnSerThrGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 609
D 2425 AACACCACCTGGAGCATGATGATACAGGCTCGCAATGATGCTGCTCAAGTTAAAGGGCT 2484
QY 610 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629
D 2485 ATAGGTTCCAGGCTTCTGTAATATCCAGATATATAGCAATCTTGGGACAGGAATAAT 2544
QY 630 IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 649
D 2545 ATATCTTCTATGTTGGCAACTAACACCTGTACTCTTAGCAATTTGACCCATCAIT 2604
QY 650 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 669
D 2605 GCAATAGCAGGAGTGTGTAATGAAATGTTCTGTCGACACAGCACTGAAAGATAAGAA 2664
QY 670 GluLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 689
D 2665 GAACTAGAAGGCTCTGGGAAGATCGGTACTGGAAGCAATAGAAAACTTCCGAACCGTTGT 2724
QY 690 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 709
D 2725 TCTTTGACTCAGAGGACAGAGTTTGAACATATATGTTGCTCAGAGTTTGCAGGTACCATAC 2784
QY 710 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 729
D 2785 AGAAACTCTTTGAGGAAAGCACACATCTTTTGAATTTACATTTTCTTCCACCAGGCAATG 2844
QY 730 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 749
D 2845 ATGTATTTTCTTGTGATGTTTCCGGTTTGGAGCCTTACTTGTGGGACATATAAACTC 2904
```

```
QY 750 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 769
D 2905 ATGAGCTTTTGGAGGATGTTCTGTATTAGTATTTTTCAGCTGTTGTCTTTGGTCCCATGGCCGTG 2964
QY 770 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 789
D 2965 GGGCAAGTCAGTTTCATTTCTGACTATGCCAAAGCCAAATATATCAGCAGGCCACATC 3024
QY 790 PheAlaLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 809
D 3025 ATCATGATCATTTGAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGCCCTAATGCTCG 3084
QY 810 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 829
D 3085 AACACATTGGAAGAAATGTCACTTTGTGTGAATTTTCAACTATATCCACCCGACCG 3144
QY 830 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 849
D 3145 GACATCCCAGTGTCTCAGGAGCTGGAGGCTGGAGGTGAAAGAGGCCAGACGCTGGGCTCTG 3204
QY 850 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 869
D 3205 GTGGCAGCAGTGGCTGTGGGAAGAGCACAGTGTCCAGCTCTCTGGAGCGGTTCTTAGCAC 3264
QY 870 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTyr 889
D 3265 CCCTTGGCAGGAAAGTGTCTGTATGTCGCAAAAGAAATAAAGCCACTGAATGTTCAGTGG 3324
QY 890 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909
D 3325 CTCGAGCAGCACCTGGGCATCTGTGCCAGAGGCCCATCTCTGTTGACTGCGACGATGTCT 3384
QY 910 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 929
D 3385 GAGACATTGCTCTATGGAGACACACAGCCGGTGGTGTACAGGAAGAGATCGTGAGGGCA 3444
QY 930 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 949
D 3445 GCAAAGGAGGCCAACATACATACATGCTTCATCGAGTCACTGCCTAATAATATAGCACTAAA 3504
QY 950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlnLysGlnArgLeuAlaIleAlaArg 969
D 3505 GTAGGAGACAAAGGAACTCAGCTCTCTGTGGCCGAAACAAACGCAATTCGCTAGCTCGT 3564
QY 970 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989
D 3565 GCCCTGTTAGACGCTCATATTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACA 3624
QY 990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009
D 3625 GAAAGTGAAGAGTTGTCCAAAGAGCCCTGGACAAAGCCAGAGAGCGCCACCTGCATT 3684
QY 1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspIleValValLeuHisAsn 1029
D 3685 GTGATTGCTCAGCGCTGTCCACCATCCAGATTCAGACTTAATAGTGTGTTCAGAAT 3744
QY 1030 GlyLysIleLysGluGlnGlyThrHisGlnLeuLeuArgAsnArgAspIleTyrPhe 1049
D 3745 GGCAGAGTCAGAGCATGGCAGCATCAGACTGCTGGCACAGAAAGGCATCTATT 3804
QY 1050 LysLeuValAsnAlaGlnSer 1056
D 3805 TCAATGCTCAGTGTCCAGGCT 3825
RESULT 12
ID ABA94366
XX ABA94366 standard; DNA; 3860 BP.
AC ABA94366;
XX
DT 26-MAR-2002 (first entry)
XX
DE Human BCRP DNA related seq Id No. 3.
```


XX Stem cell; ATP transport protein; ATP-binding cassette; antiparkinsonian;
 KW hepatotropic; neurodegenerative; cytostatic; antianemic; muscular; BCRP;
 KW cardiant; gene therapy; ds.
 XX Homo sapiens.

XX Key Location/Qualifiers
 FH 1..3843
 CDS /*tag= a

XX WO200192877-A2.

XX 06-DEC-2001.

XX 30-MAY-2001; 2001WO-US17459.

XX 31-MAY-2000; 2000US-0584586.

XX 29-MAY-2001; 2001US-086866.

XX (SUJUD-) ST JUDE CHILDREN'S RES HOSPITAL.

XX Sorrentino B, Schuetz J;

XX WPI; 2002-114368/15.

XX P-PSDB; ABB07267.

XX Identifying a stem cell, for treating e.g., muscular dystrophy,
 PT myocardial infarction, Parkinson's disease, or neurodegenerative
 PT disorders, comprises detecting the expression of an ATP transport
 PT protein (BCRP) by a cell

XX Disclosure; Page 59-60; 87pp; English.

XX The invention provides a method of identifying and/or isolating a stem
 CC cell that involves detecting the expression of an ATP transport protein
 CC containing a conserved ATP-binding cassette (BCRP) by a cell in a sample
 CC comprising stem cells. The isolated stem cells may be used in the
 CC treatment of diseases such as muscular dystrophy, degenerative liver
 CC disorder, myocardial infarction, Parkinson's disease, degenerative
 CC disorders of the brain, and for tissue regeneration or replacement.
 CC Haematopoietic cells can be used in bone marrow transplants (e.g., for
 CC treatment of leukemia) and for ex vivo gene therapy for treating blood
 CC diseases such as sickle cell anemia and thalassemia. The stem cells can
 CC also be used as cell targets in gene therapy protocols. The present
 CC sequence represents a sequence related to the BCRP for which no relevant
 CC information has been provided in the specification.

XX SQ Sequence 3860 BP; 1135 A; 746 C; 956 G; 1023 T; 0 other;

Alignment Scores:

Pred. No.: 1.88e-282 Length: 3860
 Score: 3100.50 Matches: 595
 Percent Similarity: 75.54% Conservative: 211
 Best Local Similarity: 55.76% Mismatches: 218
 Query Match: 58.17% Indels: 43
 DB: 24 Gaps: 6

US-09-873-409-4 (1-1058) x ABA94366 (1-3860)

QY 4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluVal 23

Db 712 TCATTACTGATAAGAACTCTTAGCTGTAGCTGCAAAAGCTGGAGCTAGCTGAGAGTTC 771

QY 24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGlnLysGluLeuGlnArgSer 43

Db 772 TTGCACCAATTAGAACTGTGATTGCTTGGAGGACAAAAGAAAGAACTTGAA----- 825

QY 44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63

Db 825 ----- 825

QY 64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83

Db 826 -----AGGTACACAAAAATTTAGAAAGCTAAAAAGAATTCGGATANAAG 870
 QY 84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
 Db 871 AAAGCTATTACGCCAATATTCTATAGGTGCTGCTTCTGCTGATCTATGCATCTTAT 930
 QY 104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyThrIle 123
 Db 931 GCTCTGGCCTCTCTGGTATGGGACCACCTTGGTCTCTCAGGGGAA-----TATTCTATT 984
 QY 124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
 Db 985 GGACAGTACTCACTGTATTCTTTCTGTATTAAATTTGGGGCTTTTAGTGTGGCAGGCA 1044
 QY 144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163
 Db 1045 TCTCCAAGCATTGAAGCATTTCGAATGCAAGAGGAGCAGCTTATGAATCTTCAGATA 1104
 QY 164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
 Db 1105 ATGTATAATAAGCCAGTATTGCAGCTATTGCAAGAGTGGGCACAAACACAGATAATT 1164
 QY 184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
 Db 1165 AAGGGAAATTGGAATTTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAGTTAAG 1224
 QY 204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
 Db 1225 ATCTTGAAGGGCTGAACCTGAAGTGCAGAGTGGGCACACGGTGGCCCTGGTGGAAAC 1284
 QY 224 AsnGlySerGlyLysSerThrValValGlnLeuGlnArgLeuTyrAspProAspAsp 243
 Db 1285 AGTGGCTGTGGGAGAGCACACAGTCCAGCTGATGTCAGAGGCTCTATGACCCACAGAG 1344
 QY 244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
 Db 1345 GCGATGTCAGTGTTCATGGACAGGATATTAGACCATTAATGTAAGGTTTCTACGGGAA 1404
 QY 264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283
 Db 1405 ATCATTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1464
 QY 284 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303
 Db 1465 CGCTATGGCGGTGAAATGTCCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCAAT 1524
 QY 304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323
 Db 1525 GCCTATGACTTTATCATGAAACTGCTCATAAATTTGCACCCCTGGTTGGAGAGAGAGGG 1584
 QY 324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343
 Db 1585 GCCAGTTGAGTGGTGGGAGAGAGAGAGATCGCCATTCGACGTGCCCTGGTGGTGGCAAC 1644
 QY 344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
 Db 1645 CCCAAGATCTCTGCTGATGGGCGACGTCAGCTTGGACACAGAAAGCGAGCAGTG 1704
 QY 364 ValGlnAlaAlaLeuGlnLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383
 Db 1705 GTTCAGTGGCTCTGATTAAGGCCAGAAAAGGTGCGACCACTTGTGATGATCTCATCGT 1764
 QY 384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403
 Db 1765 TTGCTACAGTTTCGTAATGCTGACGTGCTGCTGGTTCGATGATGAGTCATTTGGAG 1824
 QY 404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423
 Db 1825 AAAGGAATCATGATGAACTCATGAAAGAGAAAGGCAATTTACTTCAAACTTGTCACAATG 1884
 QY 424 Gln-----AspIleLysAlaAspGlnMetGluSerMetThr 437

Db 1885 CACACAGCAGGAATGAAGTTGAATTAGAAAAATGCAGCTGATGAATCCAAAAAGTGAAT 1944
Qy 438 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal----- 452
Db 1945 GATGCCTTGGAAATGCTCTCAATATGATTCAGATCCAGTCTAATAAGAAAAAGCAACT 2004
Qy 453 ---LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys 470
Db 2005 CGTAGAGAGTGCCTGGATCACAAGGCCAAGACAGAGAAAGCTTAGTACCAAGAGGCTCTG 2064
Qy 471 GlnLeuSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTTP 490
Db 2065 GATGAAGATATACCTCCAGTTTCTTTGGAGGATATGAAGCTAAATTTAACTGAATGG 2124
Qy 491 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 510
Db 2125 CCTATTATTTGTTGGTGTATTTTGTGCCATTATAATGGAGGCTGCACACGATTT 2184
Qy 511 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 529
Db 2185 GCAATAATATTTTCAAAGAGATTATAGGGGTTTTTACAGGAATTGATGATCTCGAAACAAA 2244
Qy 530 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 549
Db 2245 CGACAGATAGTAACCTGTTTCTACTATTGTTCTAGCCCTTGGAAATATTTCTTTTAT 2304
Qy 550 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 569
Db 2305 ACATTTTCTCTCAAGTTTCACATTGGCAAGCTGGAGAGATCTCCACCAAGCGCTC 2364
Qy 570 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 589
Db 2365 CGATACATGTTTTCGATCCATCTCAGACAGGATGTGAGTTGGTTGATGACCCCTAAA 2424
Qy 590 AsnSerThrGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 609
Db 2425 AACACCACCTGGAGCATGTACTACAGCTCGCCAATGATGCTCTCAAGTTAAAGGGCT 2484
Qy 610 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629
Db 2485 ATAGGTTCCAGGCTGCTGTAAATCCAGAAATATAGCAATCTTGGACAGGAATAT 2544
Qy 630 IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 649
Db 2545 ATATCTTCACTATGTTGGCAACTAACCTGTACTCTTAGCAATGTGACCCATCAT 2604
Qy 650 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 669
Db 2605 GCAATAGCAGGAGTTGTGAAATGAAATGTTCTCTGGCAAGCAGCTGAAAGATAGAAA 2664
Qy 670 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 689
Db 2665 GAATAGAGGTGCTGGGAAGATCGTACTGAAGCAATAGAAAACTTCCGAACCGTTGT 2724
Qy 690 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 709
Db 2725 TCTTTCACTCAGGAGCAGGAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATAC 2784
Qy 710 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 729
Db 2785 AGAAATCTTTGAGGAAGACACACATCTTTGGAAATATACATTTTCTTCCACCCAGGCAATG 2844
Qy 730 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyValaTyrLeuIleGlnAlaGlyArg 749
Db 2845 ATGATATTTTCTATGCTGATGTTCCGGTTTGGAGCCTACTTGTGGCACAATAAATC 2904
Qy 750 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 769
Db 2905 ATGAGCTTTCAGGATGTTCTGTTAGTATTTTACGCTGTTGTCGTCGTCGCGGTG 2964
Qy 770 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 789
Db 2965 GGGCAAGTCAGTTTCATTTGCTCTGCTACTATGCAAGGCCAAAATATATCAGCAGGCCCATC 3024

Qy 790 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlnGlyLysLysPro 809
Db 3025 ATCATGATCATTTGAAAAACCCCTTTGATTGACAGCTACAGCAGGAAGCCCTAATGCCG 3084
Qy 810 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 829
Db 3085 AACACATTGGAAGGAATGTCACTTTGTGTGAAGTTGTATTCAACTATCCACCCGACCG 3144
Qy 830 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 849
Db 3145 GACATCCAGTCTTCAGGACTGAGCTTGAGGTGAAAGGCCAGAGCGCTGGCTCTG 3204
Qy 850 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 869
Db 3205 GTGGCAGCAGTGGCTGTGGGAAGAGACACAGTGGTCCAGCTCTCGAGGGTGTCTACGAC 3264
Qy 870 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTTP 889
Db 3265 CCTTTGGCAGGAAGTGTCTTGTATGTCAAAGAAATAAAGCAGACTGAATGTTCACTGG 3324
Qy 890 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909
Db 3325 CTCGAGACACCTGGGCATCTGTCTCCAGGAGCCCATCTCTTTGACTGCAGCATTGCT 3384
Qy 910 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 929
Db 3385 GAGAACATTTGCTATGGAGACAACAGCCGGTGGTGTCTCACAGGAAGAGATCGTGAGGGCA 3444
Qy 930 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 949
Db 3445 GCAAGAGAGGCCAACATACATACCTTTTCATCGAGTCACTGCTTAATAAATATAGCATAA 3504
Qy 950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 969
Db 3505 GTAGGAGACAAGGAAGTCTGCTGTGGCCAGAAACACAGCAATGGCATAGTCTGT 3564
Qy 970 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989
Db 3565 GCCTTTGTTAGACAGCTCATATTTGCTTTTGGATGAAGCCAGCTCAGCTCTGGATACA 3624
Qy 990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009
Db 3625 GAAAGTGAAGAGTTTGTCCAAGAAAGCCCTGGACAAAGCCAGAGAGAGCCGACCTGCATT 3684
Qy 1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1029
Db 3685 GTGATTGCTCACCGCTGTCCACCATCCAGATGCGAGACTTAATAGTGGTGTTCAGAAT 3744
Qy 1030 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuArgAsnArgAspIleTyrPhe 1049
Db 3745 GGCAGAGTCAAGGAGCATGGCAGCATCAGCAGCTGCTGGCAGAGAAAGCATCTATTTT 3804
Qy 1050 LysLeuValAsnAlaGlnSer 1056
Db 3805 TCAATGTCAGTGTCCAGGCT 3825
RESULT 13
AARF6127
ID AARF6127 standard; cDNA; 4186 BP.
XX AARF6127;
XX 25-JUN-2001 (first entry)
DT
XX Cynomologous monkey P-glycoprotein cDNA.
DE
XX Cynomologous monkey; P-glycoprotein; PGP; multidrug transporter; MDRL;
KW efflux pump; ss.
XX
OS Macaca fascicularis.
XX
FH Key Location/Qualifiers

CDS 100...3942
 /tag= a
 /product= "PGP"
 /note= "P-glycoprotein"
 XX
 PN WO200123565-A1.
 XX
 XX 05-APR-2001.
 XX
 XX 28-SEP-2000; 2000WO-US26592.
 XX
 XX 28-SEP-1999; 99US-0156921.
 PR 12-OCT-1999; 99US-0158818.
 XX
 XX (GENT-) GENTEST CORP.
 XX
 PI Stocker PJ, Steimel-Crespi DT, Crespi CL;
 XX WPI; 2001-316136/33.
 DR P-PSDB; AAB81064.
 XX
 XX Novel isolated nucleic acid encoding cynomolgous monkey P-glycoprotein
 PT (PGP) and homologous PGP polypeptides are useful for predicting
 PT bioavailability of compound and increasing PGP transporter activity in
 PT cell -
 XX
 XX Example 1; Page 51-57; 84pp; English.
 XX
 CC This invention relates to a polynucleotide sequence encoding a
 CC cynomolgous monkey P-glycoprotein (PGP), and an allelic variant of the
 CC PGP protein. PGP, also known as multidrug transporter, MDR1 is a member
 CC of the ABC transporter superfamily. The enzyme serves as an efflux pump
 CC exporting small molecules across the cell membrane. The invention
 CC includes a cynomolgous monkey (Macaca fascicularis) PGP coding sequence
 CC and protein, and also that of an allelic variant. The PGP polynucleotide
 CC sequence is useful for increasing PGP transporter activity in a cell.
 CC Antisense sequences of the cDNA are useful for inhibiting PGP transport
 CC activity in a mammalian cell. They may also be used for increasing the
 CC bioavailability of a drug. The present sequence represents cDNA encoding
 CC cynomolgous monkey P-glycoprotein.
 XX
 SQ Sequence 4186 BP; 1226 A; 799 C; 1039 G; 1122 T; 0 other;
 Alignment Scores:
 Pred. No.: 2,12e-282 Length: 4186
 Score: 3100.50 Matches: 594
 Percent Similarity: 75.40% Conservative: 212
 Best Local Similarity: 55.57% Mismatches: 216
 Query Match: 58.17% Indels: 47
 DB: Gaps: 7
 US-09-873-409-4 (1-1058) x AAF86127 (1-4186)
 QY 4 SerLeuThrSerLysGluSerAlaTyrSerLysAlaGlyAlaValAlaGluVal 23
 Db 811 TCATTTACTGATAAAGAACTTGTAGCTTATGCAAAAGCTGGAGCAGTAGCTGAAGAGTGC 870
 QY 24 LeuSerSerLeuArgThrValIleAlaPheArgAlaGlnGlnLysGluLeuGlnArgSer 43
 Db 871 TTGGCAGCAATTAGAACTGTGATTGCTATTTGGAGGACAAAGAAAGAACTCGAA----- 924
 QY 44 PheLeuLeuAsnIleThrArgTyrAlaTyrPheTyrPheProGlnTyrPheLeuLeuSerCys 63
 Db 924 ----- 924
 QY 64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysPheGlyIleLys 83
 Db 925 -----AGGTACACACAAAAATTTAGAACAGACTAAAGAAATTTGGGATAAAG 969
 QY 84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
 Db 970 AAAGCTATTACAGCCATATTTCTATAGTGCTGCTTCTCTGCTTATCTATGATCTCTAT 1029

QY 104 GlyLeuAlaPheTrpTyrGlyThrSerLeuLeuLeuAsnGlyGluProGlyTyrThrIle 123
 Db 1030 GCTCTGGCCTTCTGGTATGGGACCACTTGGTCTCTCAAGGAA-----TATTCTATT 1083
 QY 124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
 Db 1084 GGACAAGTACTCACTGTATTCTTTCTATTAAATGGGGCTTTTAGTGTGGACAGGCA 1143
 QY 144 ValProHisPheGluThrPheAlaIleAlaArgGlyValaAlaPheHisIlePheGlnVal 163
 Db 1144 TCTCCAGCATTGAACATTGCAAAATGCAAGAGGAGCAGCTTTTGAATCTTCAAGATA 1203
 QY 164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
 Db 1204 ATTGATAATAAGCCAAAGTATTGACAGCTTATTCGAAGAGTGGGCACAAACAGATAATATT 1263
 QY 184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
 Db 1264 AAGGGAATTTGGAATTCAGAAATGTTCACTTCAGTATCCCATCTCGAAAGAAAGTTAAG 1323
 QY 204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
 Db 1324 ATCTTGAGGGCTGAACTGAAGTGCAGAGTGGCAGACGGTGGCCCTGGTGGAAAC 1383
 QY 224 AsnGlySerGlyLysSerThrValValGlnLeuGlnArgLeuTyrAspProAspAsp 243
 Db 1384 AGCGGCTGTGGGAAGAGACACACGGTCCAGCTGATGCAGAGGCTTTATGACCCACAGAG 1443
 QY 244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
 Db 1444 GCGATGCTGAGTGTGATGCGACAGGATATTAGACCATAAACGTAAGGTTTCTACCGGAA 1503
 QY 264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283
 Db 1504 ATCATCGGTGGTGGTGAAGTGCAGAACCTGTATGTTGCCACACAGATAGCTGAACAT 1563
 QY 284 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303
 Db 1564 CGCTATGGTCTGTGAAGTGTCCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 1623
 QY 304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323
 Db 1624 GCCTATGACTTTATCATGAAACTGCCTCAGAAATTTGACACCCCTGGTGGAGAGAGGG 1683
 QY 324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343
 Db 1684 GCCAGCTGAGTGTGGGAGAGAGAGAGATCGCATTCAGCTGCCCTGGTTCGCAAC 1743
 QY 344 ProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
 Db 1744 CCCAAGATCCTCTGCTGGAGAGGCCACGTCAGCTTGGACACAGAAAGTGAAGCAGTG 1803
 QY 364 ValGlnAlaLeuGlnLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383
 Db 1804 GTTCAGTGCTGCTGATGAAGGCCAGAAAGGTCGAGCACCACCATTTGTATGATGCTCATCGT 1863
 QY 384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403
 Db 1864 TTGTCTACGCTTGTATGCTGACGACGTCATCGCTGGTTCGATGATGAGTCATGTGGAG 1923
 QY 404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423
 Db 1924 AAAGGAATCATGATGAGCTCATGAAAGAGAAAGGCAATTTACTTCAAACTTGTCACAATG 1983
 QY 424 GlnAspIle-----LysLysAlaAspGlu-----GlnMet 433
 Db 1984 CAGACAGCAGGAAATGAAATTTGAATTAGAAATGACGCTGATGAATCCAAAGATGAAATT 2043
 QY 434 GluSerMetThrTyrSerThrGluArgLysThrAsnSerLeu-----Pro 448
 Db 2044 GATACCTTGAAGATGCTTTCACATGATTCAGGATCAGCTATTAAGAAAGAAAGTCCACT 2103
 QY 449 LeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThrGln 469

Db	3178	AAGCCGACAACATTGGGAAGAAATGTCA	CATTTAATGAAGTTGTATTCAACTATCCACC	3233
Qy	828	ArgProAspValPheIleLeuArgGlyLeuSerIleGluArgGlyLysThrVal	847	
Db	3238	CGACTGGACATCCAGTGTCTCAGGGCTGAGCCTGGAAGTCAAGAAGCGCCAGACGCTG	3297	
Qy	848	AlaPheValGlySerSerglyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeu	867	
Db	3298	GCCCTGTGTGGCAGCAGTGCCTGTGGAAAGACACGCTGGTCCAGCTCCTCTGGAGCGGTTC	3357	
Qy	868	TyrAspProValGlnGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnVal	887	
Db	3358	TATGACCCCTTGGCGGGAAAGTCTGTTCAGCGGAAGAAATAAAGCAACTGAATGTT	3417	
Qy	888	GlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSer	907	
Db	3418	CAGTGGCTCCAGCACAACCTGGGCATCGTGTCCAGGAGGCCATCCTCTGTTGACTGCAGC	3477	
Qy	908	IleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLys	927	
Db	3478	ATTAGTGAACAATTCCTATGGAGACACACAGCGGTGGTGTCCAGGAAGAGATCGTG	3537	
Qy	928	GluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsn	947	
Db	3538	AGGCGACCCGAAGAGGCCAATATACACGCTTCATCGAGTCACCTCCCTTAATAATATAGC	3597	
Qy	948	ThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIle	967	
Db	3598	ACCAGATTAGACACAAAGGAAGTCTCTCTGGTGGCCAGAAACAACGCGATTGCCATA	3657	
Qy	968	AlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeu	987	
Db	3658	GCTCGTCCCTTGTATAGACAGCCTCATATTTTGTCTTTTGGATGAAGCCACATCAGCTCTG	3717	
Qy	988	AspAsnAspSergluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThr	1007	
Db	3718	GATACAGAAGTGAAGAAGTTGTCCAAGAGCCCTGGCAAGCCAGAGAAGCGCGTACC	3777	
Qy	1008	CysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeu	1027	
Db	3778	TGCATTGTGATTCTCACCGCTGTCCACCATCCAGATGCGAGCTTANTAGTGGTGT	3837	
Qy	1028	HisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIle	1047	
Db	3838	CAGAATGGCAGAGTCAAGGAGCAGCGCACATCAGCAGCTGCTGGCACAGAAAGGCATC	3897	
Qy	1048	TyrPhelysLeuValAsnAlaGlnSer	1056	
Db	3898	TATTTTTCATGGTCAGTGTCCAGGCT	3924	
RESULT	14			
AD	AAF86128			
ID	AAF86128	standard; cDNA; 4195 BP.		
XX	AAF86128;			
XX	25-JUN-2001	(first entry)		
DE	Cynomolgous monkey P-glycoprotein cDNA variant 1.			
XX	Cynomolgous monkey; P-glycoprotein; PGP; multidrug transporter; MDRI;			
XX	efflux pump; ss.			
XX	Macaca fascicularis.			
XX				
FH	Key	Location/Qualifiers		
FT	CDS	100..3951		
FT		/*tag= a		
FT		/product= "PGP"		
FT		/note= "P-glycoprotein"		
FT	misc_feature	376..384		
FT		/*tag= b		
FT		/note= "Insertion of 9 nucleotides relative to PGP"		

FT . allelic variant AAF86127"

XX WO200123565-A1.

XX 05-APR-2001.

XX 28-SEP-2000; 2000WO-US26592.

XX 28-SEP-1999; 99US-0156921.

PR 12-OCT-1999; 99US-0158818.

XX (GENT-) GENTEST CORP.

XX Stocker PJ, Steimel-Crespi DT, Crespi CB;

XX WPI; 2001-316136/33.

DR P-PSDB; AAB81085.

XX Novel isolated nucleic acid encoding cynomolgus monkey P-glycoprotein (PGP) and homologous PGP polypeptides are useful for predicting bioavailability of compound and increasing PGP transporter activity in cell

XX Example 1; Page 59-65; 84pp; English.

XX This invention relates to a polynucleotide sequence encoding a cynomolgus monkey P-glycoprotein (PGP), and an allelic variant of the PGP protein. PGP, also known as multidrug transporter, MDR1 is a member of the ABC transporter superfamily. The enzyme serves as an efflux pump exporting small molecules across the cell membrane. The invention includes a cynomolgus monkey (Macaca fascicularis) PGP coding sequence and protein, and also that of an allelic variant. The PGP polynucleotide sequence is useful for increasing PGP transporter activity in a cell. Antisense sequences of the cDNA are useful for inhibiting PGP transport activity in a mammalian cell. They may also be used for increasing the bioavailability of a drug. The present sequence represents cDNA encoding cynomolgus monkey P-glycoprotein. This sequence contains a 9 nucleotide insert compared to the PGP allelic variant given in AAF86127.

SQ Sequence 4195 BP; 1230 A; 801 C; 1039 G; 1125 T; 0 other;

Alignment Scores:

Pred. No.: 2,13e-282 Length: 4195
Score: 3100.50 Matches: 594
Percent Similarity: 75.40% Conservative: 212
Best Local Similarity: 55.57% Mismatches: 216
Query Match: 58.17% Indels: 47
DB: 22 Gaps: 7

US-09-873-409-4 (1-1058) x AAF86128 (1-4195)

Qy 4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluVal 23
Db 820 TCATTTACTGATAAAGAACTCTTAGCTATGCTATGCTGAGAGGTC 879
Qy 24 LeuSerSerLeuArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43
Db 880 TTGGCAGCAATTAGAACTGCTGATTGGAGGACAAAGAAGAACTCGAA----- 933
Qy 44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63
Db 933 ----- 933
Qy 64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
Db 934 -----AGTCAACAACAAAAATTTAGAAGAAGCTAAAAAGAAATTTGGGATAAAG 978
Qy 84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
Db 979 AAAGCTATTACAGGCAATATTCTATAGGTGCTGCTTTCCTGCTTATCTATGCAATCTAT 1038
Qy 104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123
Db ----- 123

Db 1039 GCTCTGGCCTCTCTGGTATNGGGACCACTTGGTCTCTCAAGGAA-----TATTTCTATT 1092
Qy 124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
Db 1093 GCACAGTACTCACTGTATTTCTTTCTGTATTAAATTTGGGGCTTTTAGTGTGCACAGGCA 1152
Qy 144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163
Db 1153 TCTCCAAGCATTGAAGCATTTCMAATGCAAGAGGAGCAGCTTTTGAAATCTTCAAGATA 1212
Qy 164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
Db 1213 ATTGATAATAAGCCAAAGTATTGACAGCTATTCAAGAGTGGGCAAAACAGATAATAATT 1272
Qy 184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
Db 1273 AAGGGAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAGAGTTAAG 1332
Qy 204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
Db 1333 ATCTTGAAGGGCCTGAACCTGAAGGTGACAGTGGGACGCGTGGCCCTGGTTGGAAC 1392
Qy 224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243
Db 1393 AGCGGCTGTGGGAAGAGCACACGGTCCAGCTGATGACAGGCTTTATGACCCACAGAG 1452
Qy 244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
Db 1453 GGATCGTCAGTGTGATGGACAGGATATAGGACCATAAACGTAAAGTTTCTACGGGAA 1512
Qy 264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283
Db 1513 ATCATCGGTGTGTGAGTCAGGAACCTGTATTGTTGCCACCACGATAGCTGAAACATT 1572
Qy 284 LysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303
Db 1573 CGTATGCTGCTGAAGATGTCACCATGAGATGAGATTGAGAAAGCTGTCAAGAAAGCCAA 1632
Qy 304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323
Db 1633 GCTATGACTTTATCATGAAACTGCTCGCTCAGAAATTTGACACCTGTTGGAGAGAGGG 1692
Qy 324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343
Db 1693 GCCCAGCTGAGTGGTGGCAGAGAGAGGATCGCCATTGCACGTGCCCTGGTTGCAAC 1752
Qy 344 ProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
Db 1753 CCCAAGATCCTCTGCTGGACGAGGCCACGTGACCTTGGACACAGAAAGTGAAGCAGTG 1812
Qy 364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyValGlyThrThrIleValValAlaHisArg 383
Db 1813 GTTCAGTGGCTCTGGATAGGCCAGAAAGGTCGACCCACCATTTGTGATGATCTCATCGT 1872
Qy 384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403
Db 1873 TTGTCTACGGTCTGTAATGCCGACGTCTCGCTGTTTCGATGATGAGTCAITTTGGAG 1932
Qy 404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423
Db 1933 AAAGGAATCATGATGAGCTCATGAAAGAGAAAGGAGGATTTACTTCAAACTTGTCACAATG 1992
Qy 424 GlnAspIle-----LysLysAlaAspGlu-----GlnMet 433
Db 1993 CAGACAGCAGGAATGAAATTAATTAAGAAATGCGAGCTGATGATCAAAAGAGTGAATT 2052
Qy 434 GluSerMetThrTyrSerThrGluArgLysThrAsnSerLeu-----Pro 448
Db 2053 GATACCTTGGAAATGCTTCACATGATTCAGGATCCAGTCTAATAAGAAAAAGATCCACT 2112
Qy 449 LeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThrGln 468
Db 2113 CGTAGGAGTGTCTGATCAAGGCCAA-----GACAGAAAGCTTAGTACCAAGAG 2166


```
QY 469 SerLysGluLeuSerLeuProGluValSerLeuLeuLysLeuLysLeuLysPro 488
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 2167 GCTCTGGATCAAAAGTATACCTCCAGTTCTCTTTGGAGGATTATGAAGCTAAATTAACT 2226

QY 489 GluTrpProPheValValLeuGluThrLeuAlaSerValLeuAsnGlyThrValHisPro 508
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 2227 GAGTGGCCCTTATTTTGTGTGTATTTTGGCCATTATAAATGGAGTCTCCAAACCA 2286

QY 509 ValPheSerIlePheAlaLysIleIleThrMetPhe--GlyAsnAsnAspLysThr 527
   ||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 2287 GCATTTGCGAGTATATTTCAAGAGATTATAGGATTTTACAGAAGATGATGATGCCGAA 2346

QY 528 ThrLeuLysHisAspAlaGluIleThrSerMetIlePheValIleLeuGlyValIleCys 547
   ||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 2347 ACAAAACGACAGAAATAGTAACCTTGTCTTCACTATTGTTCTAGCTCTTGAATGTCTT 2406

QY 548 PheValSerThrPheMetGlnGlyLeuPheThrGlyArgAlaGlyGluIleLeuThrMet 567
   ||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 2407 TTTATTACATTTTCTTCAGGGCTTCACATTTGGCAAAAGCTGGAGAGATCCCTCACCAAG 2466

QY 568 ArgLeuArgHisLeuAlaPheLysAlaMetLeuLysGlnAspIleAlaTrpPheAspGlu 587
   ||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 2467 CGGCTCCGATACATGTTTCCGATCATGCTCAGACAGGATGTGAGCTGGTTGATGAC 2526

QY 588 LysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGln 607
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 2527 CCTAAAACACCACTGGAGCATTTGACTACAGGCTCGCCCAATGATGCTGCTCAAGTTAA 2586

QY 608 GlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSer 627
   ||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 2587 GGGGCTATAGTTTCCAGGCTGTCTATATATACCCAGAATATAGCAAAATCTTGGGACAGGA 2646

QY 628 ValIleIleSerPheIleThrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaPro 647
   ||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 2647 ATAATTATATCTTAACTATGTTTGGCACTGCACTGTTACTCTTAGCAATTGTACCC 2706

QY 648 ValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAsp 667
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 2707 ATCATTTGCAATAGCAGGAGTTGTTGAAATGAAATGTTGCTCGACACGACCTGAAAGAT 2766

QY 668 LysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThr 687
   ||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 2767 AAGAAAGAACTAGAGAGTGTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAAC 2826

QY 688 IleValSerLeuThrArgGluLysAlaPheGluGluMetThrGluGluMetLeuGlnThr 707
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 2827 GTTGTTTCTTTGACTCAGGAGCGAAGTTTGAACATATGATGATCAGAGTTTGCAGGTA 2886

QY 708 GlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysThrAlaPheSerHis 727
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 2887 CCATACAGAAACTCTTTGAGGAAGGCACACATCTTTGGAATCATCGTTTCTTCACGCGAG 2946

QY 728 AlaPheIleThrPheAlaThrAlaAlaGlyPheArgPheGlyAlaThrIleGlnAla 747
   ||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 2947 GCAATGATGATATTTTCTTCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGTGGGCACAT 3006

QY 748 GlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaThrGlyAlaMet 767
   ||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 3007 AGTCTCATGAGCTTTGAGGATGTTCTGTAGTATTTTTCAGCTGTTGCTTGTGGTGCATG 3066

QY 768 AlaIleGlyLysThrLeuValLeuAlaProGluThrSerLysAlaLysSerGlyAlaAla 787
   ||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 3067 GCCGTGGGCAAGTCAGTTCTATTTGCTCTGACTATGCCAAAGCCAAAGTATCAGCAGCC 3126

QY 788 HisLeuPheAlaLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLys 807
   ||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 3127 CACATCATCATGATCATTTGAAAAAACCCTTTTGTGATGACAGCTACAGCAGAGAGCCTA 3186

QY 808 LysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPheThrProCys 827
   ||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 3187 AAGCCGAACACATTTGGAAGGAATGTACACATTTAATGAAGTTGTATTTCACTATCCACC 3246
```

```
QY 828 ArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrVal 847
   ||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 3247 CGACTGGACATCCCACTGCTTCAGGGCTGAGCTCGAAAGTGAAGAGGCCAGACGCTG 3306

QY 848 AlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeu 867
   ||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 3307 GCCCTGTTGGGAGCAGTGGCTGTGGAAAGACGCGTGGTCCAGCTCTCTGGAGCGGTTTC 3366

QY 868 TyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnVal 887
   ||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 3367 TATGACCCCTTGGCGGGAAAGTGTCTGTGACGCAAAAGAAATAAAGCAACTGATGTT 3426

QY 888 GlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSer 907
   ||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 3427 CAGTGGCTCCGAGCACACCTGGGCATCGTGTCCAGGAGCCCATCTCTGTTGACTGCAGC 3486

QY 908 IleAlaGluAsnIleAlaThrGlyAspAsnSerArgValValProLeuAspGluIleLys 927
   ||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 3487 ATTAGTGAGAACTATGCTTATGGAGACAACAGCCGGTGGTGTACAGGAAGAGATCGTG 3546

QY 928 GluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysThrAsn 947
   ||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 3547 AGGCGACGCCAAGAGGCCAATATACACGCTTCATCGAGTCACTGCTTAATAATATAGC 3606

QY 948 ThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIle 967
   ||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 3607 ACCAGAGTAGGAGACAAAGAACTCAGCTCTCTGGTGGCCAGAAAACAACGCATTGGCATA 3666

QY 968 AlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeu 987
   ||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 3667 GCTCGTGCCTTGTAGACAGCCTCATATTTTGTCTTTCGATGAAAGCCACATCAGCTCTG 3726

QY 988 AspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThr 1007
   ||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 3727 GATACAGAAAGTGAAGAGTTGTCCAAGAGCCCTGGACAAAGCCAGAGAGCCCTACC 3786

QY 1008 CysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeu 1027
   ||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 3787 TGCATTGTGATTGCTCACCGCTGTCCACCATCCAGATCCAGACTTAATAGTGTGTTT 3846

QY 1028 HisAsnGlyLysIleLysGluGlnGlyThrHisGlnLeuLeuLeuArgAsnArgAspIle 1047
   ||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 3847 CAGAATGGCAGAGTCAAGGAGCACGCGCACATCAGCAGCTGCTGGCACAGAAAGCATC 3906

QY 1048 TyrPheLysLeuValAsnAlaGlnSer 1056
   ||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 3907 TATTTTCAATGTCAGTGTCCAGGCT 3933

RESULT 15
AAV6533
ID AAV6533 standard; cDNA; 4264 BP.
XX
AC AAV6533;
XX
DT 08-JAN-1999 (first entry)
XX
DE Mutated human P-glycoprotein gene designated mdr-delta-F335.
XX
KW mdr1-delta-F335; cyclosporin-resistant deletion mutant; human;
KW P-glycoprotein; cyclosporin; drug efflux;
KW modulator-resistant multidrug resistance; cancer therapy; ss.
OS Homo sapiens.
XX
PN US5830697-A.
XX
PD 03-NOV-1998.
XX
PF 21-JAN-1997; 97US-0784649.
XX
PR 21-JAN-1997; 97US-0784649.
XX
PA (STRD ) UNIV LELAND STANFORD JUNIOR.
```


XX Chen G, Sikic BI;
 PI WPI; 1998-609244/51.
 XX
 XX DNA encoding mutant P-glycoprotein polypeptide - useful for
 PT producing recombinant polypeptide
 XX
 XX Disclosure; Columns 15-20; 17pp; English.
 PS
 CC The present sequence represents mdrl-delta-F335, a gene encoding a
 CC cyclosporin-resistant deletion mutant of human P-glycoprotein lacking
 CC Cys335. The recombinant protein can be used to produce antibodies. It
 CC can also be used in drug screening assays to study the effects of
 CC modulators such as cyclosporin on P-glycoprotein-mediated drug efflux.
 CC The nucleic acid can be used to transfer modulator-resistant multidrug
 CC resistance into transfected cells, e.g. haematopoietic stem cells for
 CC use during cancer therapy.
 XX
 SQ Sequence 4264 BP; 1263 A; 807 C; 1041 G; 1153 T; 0 other;
 Alignment Scores:
 Pred. No.: 2,18e-282 Length: 4264
 Score: 3100.50 Matches: 595
 Percent Similarity: 75.54% Conservative: 211
 Best Local Similarity: 55.76% Mismatches: 218
 Query Match: 58.17% Indels: 43
 DB: 19 Gaps: 6
 US-09-873-409-4 (1-1058) x AAV66533 (1-4264)
 QY 4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValaGluVal 23
 DB TCATTACTGATAAAGAACTCTTAGCTGTATGCAAAAGCTGGAGCAGTAGCTGAAGGTC 909
 QY 24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43
 DB 910 TTGCGCAAAATTAGAACTGTGATTCATTCATTTGGAGGACAAAAGAAAGAACTTGAA- 963
 QY 44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63
 DB 963 ----- 963
 QY 64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
 DB -----AGGTACACAAAATTAGAAAGCTTAAAGAACTTGGATTAAG 1008
 QY 84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
 DB 1009 AAAGCTATTACAGCAATATTCTATAGTGTCTCTTCCTGCTGTATGATCTATGATCTTAT 1068
 QY 104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123
 DB 1069 GCTCTGGCCTTCTGTATGGGACCACCTTGTCTCTCAGGGGAA-----TATTATT 1122
 QY 124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
 DB 1123 GGCAAGTACTCACTGATTTCTTCTGTATTAATGGGGCTTTAGTGTGGACGGCA 1182
 QY 144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163
 DB 1183 TCTCCAAGCATTTGAAGCATTTTGAATGCAAGAGGACGAGCTTATGAATCTTCAAGATA 1242
 QY 164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
 DB 1243 ATTGATAATAAGCCCAAGTATTGACAGCTATTGCAAGAGTGGGCAACCAACAGATAAATT 1302
 QY 184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
 DB 1303 AAGGGAATTGGAAATTGAGAAATGTTCACTTCAGTTACCAATCTCGAAAGAGTTAAG 1362
 QY 204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223

DB 1363 ATCTTGAAGGGCTGAACCTGAAGGTGCAGAGTGGGCGACAGCGTGGTGGTGAAC 1422
 QY 224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243
 DB 1423 AGTGGCTGTGGGAGAGACACACAGTCCAGCTGATCCAGAGGCTCTATGACCCACAGAG 1482
 QY 244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
 DB 1483 GGGATGTCAGTGTTCATGACACAGGATATTAGACCATAAATGTAAGGTTTCTACGGAA 1542
 QY 264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283
 DB 1543 ATCATTTGGTGTGGTGTAGTCAGGAACCTGTATTGTTGCCACCCAGTAGCTGAAACATT 1602
 QY 284 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303
 DB 1603 CGCTATGGCGGTGAAATGTCCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAT 1662
 QY 304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323
 DB 1663 GCCTATGACTTTTATCATGAACCTGCCTCATAAATTTGACACCCCTGTTGGAGAGAGGG 1722
 QY 324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343
 DB 1723 GCCCAGTTGAGTGTGGGAGAGCAGAGAGATGCCCATTTGCACGTCCCTGGTTCGCAAC 1782
 QY 344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
 DB 1783 CCCAAGATCTCTGCTGTGATGAGCCAGCTGACGCTTGGACACAGAAAGCAGAGAGTG 1842
 QY 364 ValGlnAlaAlaLeuLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383
 DB 1843 GTTCAGTGTGCTCTGATTAAGGCCAGAAAAGGTGCGACCACTTGTGTAGTCTCATCGT 1902
 QY 384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403
 DB 1903 TTGCTACAGTTTCGTAATGCTGACGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1962
 QY 404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423
 DB 1963 AAAGGAATCATGATGAACTCATGAAAGAGAAAGGCAATTACTTCAAACTTGTGCACATG 2022
 QY 424 Gln-----AspIleLysLysAlaAspGluGlnMetGluSerMetThr 437
 DB 2023 CAGACAGCAGGAATAAAGTTGAATTAGAAAATGCGAGCTGATGAATCCAAAAGTGAAT 2082
 QY 438 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal----- 452
 DB 2083 GATGCTTGGAAATGCTTCAATGATTCAAGATCCAGTCTAATAAGAAAAGATCAACT 2142
 QY 453 ---LysSerIleLys---SerAspPheIleAspLysAlaGluSerThrGlnSerLys 470
 DB 2143 CGTAGAGTGTGCTGGATCAAGCCCAAGCAGAAAGCTTAGTACCAAGAGGCTCTG 2202
 QY 471 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrp 490
 DB 2203 GATGAAGTATATCTCCAGTTTCTTTGGAGGATTAAGCTAAATTTAACTGAATGG 2262
 QY 491 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 510
 DB 2263 CCTTATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2322
 QY 511 SerIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 529
 DB 2323 GCAATATATTTTCAAGATTAAGGGTTTTTACAGAAATGATGATCTGAAACAAA 2382
 QY 530 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 549
 DB 2383 CGACAGAAATAGTAACCTGTTTTCACATTTGTTTCTAGCCCTTGAATATTCTTTTATT 2442
 QY 550 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 569
 DB 2443 ACATTTTCTCAGGGTTTTCACATTTGCAATTTGGCAAGCTGGAGAGATCTCCACCAAGCGGCTC 2502


```
QY 570 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 589
DB 2503 CGATACATGTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTGATGACCCCTAAA 2562
QY 590 AspSerThrGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 609
DB 2563 AACACCACTGGAGCATTTGATACAGGCTCGCAATGATGCTCAAGTTAAAGGGGCT 2622
QY 610 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629
DB 2623 ATAGGTTCCAGGCTGCTGTAATTTACCCAGATATAGCAATCTTGGCAGAGNAATTT 2682
QY 630 IleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeu 649
DB 2683 ATATCTCTCATCTATGTTGGCAACTAACACTGTTACTCTTAGCAATTTGACCCATCAT 2742
QY 650 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 669
DB 2743 GCAATAGCAGGAGTTGTTGAATGAAATGTTGCTGGCAAGCACCTGAAAGATTAAGAAA 2802
QY 670 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 689
DB 2803 GAACAGAGGCTGCTGGGAAGATCGCTACTGAGCAATAGAAAACCTTCCGAAACCGTTGTT 2862
QY 690 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 709
DB 2863 TCCTTGACTCAGGAGCAGAAAGTTTGAACATATGATGCTCAGAGTTTTCAGGTTACCATC 2922
QY 710 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 729
DB 2923 AGAAACTCTTTGAGGAAAGCACACATCTTTGGAAATTACATTTTCTTCAACCCAGGCAATG 2982
QY 730 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 749
DB 2983 ATGTATTTTTCCTATGCTGATGTTCCGGTTTGGAGCCTACTTGTGGGCACATAAACTC 3042
QY 750 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 769
DB 3043 ATGAGCTTTCAGGAGTGTCTGTTAGTATTTCAGCTGTTGCTTGTGGTCCATGGCCGTG 3102
QY 770 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 789
DB 3103 GGGCAAGTCAGTTTCATTTGCTGCTGACTATGCCAAAGCCAAAATATCAGCAGGCCACATC 3162
QY 790 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 809
DB 3163 ATCATGATCAITGAAAAAACCCCTTTGATTGACAGCTACAGCAGGGAAGCCCTAATGCCG 3222
QY 810 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPheTyrProCysArgPro 829
DB 3223 AACACATTGGAGGAATGTACACATTTGGTGAAGTTGTATTCAACTATCCACCCGACCG 3282
QY 830 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 849
DB 3283 GACATCCAGTGTCTTCAGGGACTGAGCCTGGAGGTGAAGAGGCCAGACGCTGGCTCTG 3342
QY 850 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 869
DB 3343 GTGGCAGCAGTGGCTGTGGGAAGAGCACAGTGTCTCAGCTCTCGAGCGGTTCACGAC 3402
QY 870 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 889
DB 3403 CCCTTGGCAGGGAAGTGTCTGTGATGGCAAGAAATAAGCGACTGATTTTCAGTGG 3462
QY 890 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909
DB 3463 CTCGAGCACACCTGGGCATCGTCTCCAGAGGCCATCTCTGTTTACTGTCAGCATTTGCT 3522
QY 910 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 929
DB 3523 GAGAACATTGCCTATGGAGACACAGCCGGTGGTGTCCAGGAAGAGATTGTGAGGGCA 3582
```

```
QY 930 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 949
DB 3583 GCAAGGAGGCCCAACATACATATGCTTTCATCGAGTCACTGCCTAATAAATATAGCACTAAA 3642
QY 950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 969
DB 3643 GTAGGAGACAAAGAACTCAGCTCTCTGTGGCCAGAAACAAACGCAATGGCCATGCTCGT 3702
QY 970 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989
DB 3703 GCCCTTGTAGACAGCCTCATATTTTGTGTTGGATGAAGCCAGCTCAGCTCTGGATACA 3762
QY 990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009
DB 3763 GAAAGTGAAGAGTTTGTCCAAGAAAGCCCTGGACAAAGCCAGAGAGGCCGACCTGCATT 3822
QY 1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1029
DB 3823 GTGATTGCTCACCGCTGCTCCACCATTCCAGATGCAGACTTAATAGTGTGTTCAGAAAT 3882
QY 1030 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1049
DB 3883 GGCAGAGTCAGGAGCATGGCAGCATCAGCAGCTTAAATAGTGTGTTCAGAAAT 3942
QY 1050 LysLeuValAsnAlaGlnSer 1056
DB 3943 TCAATGGTCAGTGTCCAGGCT 3963
```

Search completed: March 30, 2003, 03:12:46
Job time : 1302.02 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 31, 2003, 02:24:20 ; Search time 154.939 Seconds
(without alignments)
2094.145 Million cell updates/sec

Title: US-09-873-409-4

Perfect score: 5330

Sequence: 1 MVISLTSKLSAYSAGAVA.....QELLNRDIYFKLVNAQSVQ 1058

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -DEV=xlh

-MODEL=frame.p2n.model -DEV=xlh
-Q/cgn2_1/USFTO_spool/US09873409/runat_27032003_115422_19259/app_query.fasta_1.7544
-DB=issued Patents NA -QFMT=fastap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09873409 @CGN 1 1 578 @runat_27032003_115422_19259 -NCPUS=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:
1: /cgn2_6/prodata/2/ina/5A COMB.seq:*
2: /cgn2_6/prodata/2/ina/5B COMB.seq:*
3: /cgn2_6/prodata/2/ina/6A COMB.seq:*
4: /cgn2_6/prodata/2/ina/6B COMB.seq:*
5: /cgn2_6/prodata/2/ina/PCTUS COMB.seq:*
6: /cgn2_6/prodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3100.5	58.2	4264	2	US-08-784-649A-1
2	3100.5	58.2	4264	2	US-08-784-649A-1
3	3100.5	58.2	4646	1	US-08-181-471-2
4	3100.5	58.2	4669	6	5206352-3
5	3100.5	58.2	6505	2	US-08-793-610-5
6	3100.5	58.2	9318	2	US-08-793-610-6
7	3086	57.9	4233	3	US-09-120-513-1
8	3086	57.9	4233	4	US-09-450-105-1
9	3073.5	57.7	4669	2	US-08-752-447-1
10	3073.5	57.7	4669	4	US-09-316-167-1
11	3056.5	57.3	4669	2	US-08-583-276-18
12	2280	42.8	2726	1	US-08-461-823-1

13	2067	38.8	4002	2	US-08-996-545-1	Sequence 1, Appli
14	2067	38.8	4002	2	US-08-996-545-3	Sequence 3, Appli
15	2067	38.8	4002	4	US-09-328-320-1	Sequence 1, Appli
16	2067	38.8	4002	4	US-09-328-320-3	Sequence 3, Appli
17	2038.5	38.2	4047	2	US-08-612-734B-1	Sequence 1, Appli
18	1995.5	37.4	4800	2	US-08-612-734B-3	Sequence 3, Appli
19	1926	36.1	3924	1	US-08-993-246C-1	Sequence 1, Appli
20	1885.5	35.4	4224	1	US-08-612-521-1	Sequence 1, Appli
21	1558	29.2	6143	1	US-08-612-521-3	Sequence 3, Appli
22	1546	29.0	3909	1	US-08-232-537-1	Sequence 1, Appli
23	1544.5	29.0	3924	2	US-08-996-644-3	Sequence 3, Appli
24	1544.5	29.0	3924	2	US-09-352-552-3	Sequence 3, Appli
25	1544.5	29.0	3927	2	US-08-996-644-1	Sequence 1, Appli
26	1544.5	29.0	3927	3	US-09-352-552-1	Sequence 1, Appli
27	1460.5	27.4	3792	4	US-09-351-224E-10	Sequence 10, Appli
28	1326	24.9	3999	4	US-09-351-224E-9	Sequence 9, Appli
29	1037	19.5	13188	4	US-08-961-527-70	Sequence 70, Appli
30	982.5	18.4	7760	4	US-08-961-527-63	Sequence 63, Appli
31	936.5	17.6	4403765	4	US-09-103-840A-2	Sequence 2, Appli
32	936.5	17.6	4411529	4	US-09-103-840A-1	Sequence 1, Appli
33	935.5	17.6	6492	4	US-08-961-527-188	Sequence 188, App
34	898	16.8	7186	4	US-08-961-527-39	Sequence 39, Appli
35	678	12.7	2376	1	US-08-394-880B-1	Sequence 1, Appli
36	659.5	12.4	5889	1	US-08-463-092B-5	Sequence 5, Appli
37	659.5	12.4	5889	2	US-08-462-109A-5	Sequence 5, Appli
38	659.5	12.4	5889	2	US-08-460-907B-5	Sequence 5, Appli
39	659.5	12.4	5889	3	US-08-463-179A-5	Sequence 5, Appli
40	659.5	12.4	5889	3	US-08-461-384B-5	Sequence 5, Appli
41	657	12.3	4781	2	US-09-001-273-1	Sequence 1, Appli
42	657	12.3	4781	4	US-08-843-459A-1	Sequence 1, Appli
43	657	12.3	4847	3	US-09-061-400-1	Sequence 1, Appli
44	649	12.2	5011	1	US-08-463-092B-3	Sequence 3, Appli
45	649	12.2	5011	2	US-08-462-109A-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-08-784-649A-1
; Sequence 1, Application US/08784649A
; Patent No. 5830697
; GENERAL INFORMATION:
; APPLICANT: Sikic, Branimir I
; TITLE OF INVENTION: P-GLYCOPROTEIN MUTANT RESISTANT TO
CYCLOSPORIN MODULATION
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 2200 Sand Hill Road
; CITY: Menlo Park
; STATE: CA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/784,649A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: Reg. No. 5830697 36,677
; REFERENCE/DOCKET NUMBER: 06037/007001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-322-5070
; TELEFAX: 415-854-0875
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4264 base pairs

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-784-649A-1

Alignment Scores:

Pred. No.: 0
Score: 3100.50
Percent Similarity: 75.54%
Best Local Similarity: 55.76%
Query Match: 58.17%
DB: 2
Gaps: 6

US-09-873-409-4 (1-1058) x US-08-784-649A-1 (1-4264)

QY 4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluVal 23
DB 850 TCATTCTGATAAAGAACTCTTAGCGTATGCAAAAGCTGGAGCAGTAGCTGAAGAGTC 909
QY 24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43
DB 910 TTGGCAGCAATTAGAACTGTGATTCATTTGGAGGACAAAGAAAGAACTTGAA----- 963
QY 44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63
DB 963 ----- 963
QY 64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
DB 964 -----AGGTACACAAATAATTTAGAAAGCTAAAGAACTTTGGGATAAG 1008
QY 84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
DB 1009 AAAGCTATTACAGCAATAATTTCTAGGTGCTCTCTTCCTGCTGATGCTATGATCTTAT 1068
QY 104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123
DB 1069 GCTCTGGCCCTCTGATGGGACCACTTGCTCTCTCTCAGGGGAA-----TATCTATT 1122
QY 124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
DB 1123 GGACAAGTACTCACTGATTCTCTTCTGTATTAATTTGGGCTTTTAGTTGGACAGCA 1182
QY 144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaPheHisIlePheGlnVal 163
DB 1183 TCTCAAGCATTTGAAGCATTTTGAATTCGAAGAGCAGCTTATGAAATCTTCAAGATA 1242
QY 164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
DB 1243 ATTGATAAAGCCCAAGTATTGACAGCTATTGGAAGAGTGGGCACAAACACAGATAATT 1302
QY 184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
DB 1303 AAGGGAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAGAAAGTTAAG 1362
QY 204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
DB 1363 ATCTTGAAGGCCCTCGAACTGAAGCTGAGAGTGGGAGCAGCGTGGCCCTGTTGGAAC 1422
QY 224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243
DB 1423 AGTGGCTGTGGAGAGACACACAGTCCAGCTGATGACAGGCTCTATGACCCACAGAG 1482
QY 244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
DB 1483 GGGATGGTCACTGTTGATGGACAGGATATTAGGACCAATAAATGTAAGGTTTCTACGGGA 1542
QY 264 HistLeGlyValValSerGlnProValLeuPheGlyThrThrIleSerAsnAsnIle 283
DB 1543 ATCAATGGTGGGTGGAGTCAGAACTGTTATGTTGCCACCAAGTAGCTGAAACATT 1602
QY 284 LysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303

DB 1603 CGCTATGGCGGTGAAATGTCCACATGGATGAGATTGAGAAAGCTGTCAAGGAAGCAAT 1662
QY 304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323
DB 1663 GCCTATGACTTTATCATGAACCTGCTCATAAATTTGACACCCCTGGTTGGAGAGAGGG 1722
QY 324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343
DB 1723 GCCCAGTTGAGTGGCGGAGAGCAGAGATCGCCATTGCACGTGCCCTGGTTCCGAAC 1782
QY 344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
DB 1783 CCCAAGATCTCTCTGCTGATGAGGCCACGTGAGCCCTTGGACACAGAAAGCGAAGCAGTG 1842
QY 364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValAlaAlaHisArg 383
DB 1843 GTTCAGGTGCTCTGATGAAGCCAGAAAGGTCGACACCACTTGTATGATGCTCATCGT 1902
QY 384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403
DB 1903 TTGCTACAGTTCGTATGCTGACGTCACTCGCTGTTTCGATGATGAGTCAATTGTGGAG 1962
QY 404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423
DB 1963 AAAGGAAATCATGATGAATCATGAAAGAGAAAGGCACTTACTTCAAACTTGTCACAATG 2022
QY 424 Gln-----AspIleLysLysAlaAspGluGlnMetGluSerMetThr 437
DB 2023 CAGACAGCAGGAAATGAAGTTGAATTAGAAATGACGCTGATGAATCCAAAGTAGAAATT 2082
QY 438 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal----- 452
DB 2083 GATGCTTGGAAATGCTTCAATGATTCAGATCCAGTCACTATAGAAAAAGATCAACT 2142
QY 453 ---LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys 470
DB 2143 CGTAGGAGTGTCCGTGGATCACAAGCCCAAGCAGACAGAAAGCTTAGTACCAAGAGGCTCTG 2202
QY 471 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuLysProGluTrp 490
DB 2203 GATGAAGTATACCTCCAGTTCTCTTTGGAGGATATGAACTAAATTTAACTGAATGG 2262
QY 491 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 510
DB 2263 CCTATTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2322
QY 511 SerIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 529
DB 2323 GCAATAATATTTTCAAGATATATAGGGTTTTTACAGAATTTGATGATCCCTGAACAAAA 2382
QY 530 LysHisAspAlaGluLeuTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 549
DB 2383 CGACAGATAGTAACTGTTTTCACCTATTGTTTTCAGCCCTGGAATTTATTTCTTTATT 2442
QY 550 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 569
DB 2443 ACATTTTCTTCAGGGTTTCATTTTGGCAAGCTGGAGAGATCTCTCACCAGCGGCTC 2502
QY 570 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 589
DB 2503 CGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTGATGACCCCTAAA 2562
QY 590 AsnSerThrGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 609
DB 2563 AACACACCTGGAGCATGATGATCCAGGCTCGCAATGATGCTCTCAAGTTAAAGGGGCT 2622
QY 610 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629
DB 2623 ATAGTTTCCAGGCTTCTGTTAATACCAGAAATATAGCAATCTTGGGACAGGAATAATT 2682
QY 630 IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 649

Db 2683 ATATCTCTCATCTATGTTGGCACTAAACACTGTTACTCTTAGCAATTTGACCCATCAT 2742
Qy 650 AlaValThrGlyMetLeuGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 669
Db 2743 GCAATAGCAGGAGTGTGTGAATGAAATGTTGCTGCGACAGCACTGAAAGATAAGAA 2802
Qy 670 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnLysArgThrIleVal 689
Db 2803 GAAC TAGAAGGTCTGGGAAGATCGCTACTGAAGCAATAGAAAACCTTCCGAACCGTGT 2862
Qy 690 SerLeuThrArgGluLysAlaPheGluGlnMetThrGluGluMetLysGlnThrGlnHis 709
Db 2863 TCCTTACTCAGAGCAGAGGAGTTTGAACTATGATGTGCTCAGAGTTTGAGGTACCATAC 2922
Qy 710 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 729
Db 2923 AGAACTCTTTGAGGAAGCAGACATCTTTTGAATATACATTTTCTTCCACCCAGGCAATG 2982
Qy 730 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrIleuLeuGlnAlaGlyArg 749
Db 2983 ATGATATTTTCCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCAGCAATAA 3042
Qy 750 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 769
Db 3043 ATGAGCTTGGAGATGTTCTGTAGTATTTTACGTGTTCTTGGTGGCAGTGGCGGTG 3102
Qy 770 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 789
Db 3103 GGGCAAGTCAGTTCACTTGTCTCTGACTATGCCAAGCCAAATATATCAGCAGCCACATC 3162
Qy 790 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlnGlyLysPro 809
Db 3163 ATCATGATCATTTGAAAGAACCCCTTTGATTGACAGCTACAGCAGGAGCCCTTAATGCCG 3222
Qy 810 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 829
Db 3223 AACACATTGGAAGAAATGTACATTTGGTGAATGTGTTCACTATCCACCCGACCG 3282
Qy 830 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 849
Db 3283 GACATCCAGTGTCTCAGGGACTGAGCTGGAGGTGAAGAGGCCAGAGCGCTGGCTCTG 3342
Qy 850 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 869
Db 3343 GTGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGAC 3402
Qy 870 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 889
Db 3403 CCTTGGCAGGGAAGTGGCTGTGATGGCAAGAAATAAGGAGCTGAATGTTCACTGG 3462
Qy 890 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909
Db 3463 CTCGAGCACACCTGGGCATCGTGTCCAGAGGCCATCTGTTGACTGCAGCATGCT 3522
Qy 910 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 929
Db 3523 GAGAAATTCCTATGAGACAAACAGCGCGGTGGTGTACAGGAAGAGATGTTGAGGCGCA 3582
Qy 930 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 949
Db 3583 GCAAAGGAGGCCAACATACATGCTTATCGAGTCACTGCTTAATAATATAGCACTAAA 3642
Qy 950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 969
Db 3643 GTAGGACAAAGAACTCAGCTCTGTGGTGGCCAGAAACAGCGATGTCATGCTCGT 3702
Qy 970 AlaLeuLeuGlnProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989
Db 3703 GCCCTTGTAGACAGCCTCATATTTGCTTTTGGATGAGCCAGCTCAGCTCTGGATACA 3762
Qy 990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009
Db 3763 GAAAGTGAAAGGTTGTCCAAAGAGCCCTGGACAAAGCCAGAGAGGCCGACCTGCATT 3822

Qy 1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspIleuLeuValValLeuHisAsn 1029
Db 3823 GTGATTGCTCAGCGCTGTCACCATCCAGATTCAGACTTAATAGTGGTGTTCAGAA 3882
Qy 1030 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1049
Db 3883 GGCAGAGTCAAGAGCATGCGCAGCATCAGACGCTCTGGCACAGAAGGCATCTATTT 3942
Qy 1050 LysLeuValAsnAlaGlnSer 1056
Db 3943 TCAATGTCAGTGTCCAGGCT 3963
RESULT 2
US-08-784-649A-5
; Sequence 5, Application US/08784649A
; Patent No. 5830697
; GENERAL INFORMATION:
; APPLICANT: Sikic, Branimir I
; TITLE OF INVENTION: P-GLYCOPROTEIN MUTANT RESISTANT TO
; TITLE OF INVENTION: CYCLOSPORIN MODULATION
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 2200 Sand Hill Road
; CITY: Menlo Park
; STATE: CA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/784,649A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: Reg. No. 5830697 36,677
; REFERENCE/DOCKET NUMBER: 06037/007001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-322-5070
; TELEFAX: 415-854-0875
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4264 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-784-649A-5
Alignment Scores:
Pred. No.: 0 Length: 4264
Score: 3100.50 Matches: 595
Percent Similarity: 75.54% Conservative: 211
Best Local Similarity: 55.76% Mismatches: 218
Query Match: 58.17% Indels: 43
DB: 2 Gaps: 6

US-09-873-409-4 (1-1058) x US-08-784-649A-5 (1-4264)
Qy 4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluVal 23
Db 850 TCATTTACTGATAAAGAACTCTTAGCGGTATGCAAAAGCTGGAGCAGTAGCTGAAGAGGTC 909
Qy 24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43
Db 910 TTGGCAGCAATTAGAACTGTGATTGCTTTGGAGGACAAAGAAAGAACTTGAA----- 963

QY 44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63
Db 963 ----- 963
QY 64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspPheGlyIleLys 83
Db 964 -----AGGTACACAAATAATTTAGAAAGCTAAAGAAATGGGATAAG 1008
QY 84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyTyr 103
Db 1009 AAGCTATTACAGCAATATTTCTAGGTGCTGCTTCTGCTGATCTATGATCTAT 1068
QY 104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleAsnGlyGluProGlyTyrThrIle 123
Db 1069 GCTCTGSCCTTCTGTTATGGACACCTTGGTCTCTCAGGGAA-----TATTCTATT 1122
QY 124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
Db 1123 GGACAAGTACTCACTGTATTCTTTCTGTATTAAATTGGGCTTTAGTGTGGACAGGCA 1182
QY 144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163
Db 1183 TCTCCAAAGCAATTGAAGCATTTGCAATTCGAAGAGTGGGCACAAACAGATAATATT 1242
QY 164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
Db 1243 ATTGATAATAAGCCAACTATTGACAGCTATTTCGAAGAGTGGGCACAAACAGATAATATT 1302
QY 184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
Db 1303 AAGGGAAATTTGGAATTCAGAAATGTTTCACTTCAGTTTACCCTCTCGAAAGAAAGTTAAG 1362
QY 204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
Db 1363 ATCTTGAGGCGCTGAACCTGAGGTGCAGAGTGGCAGCGTGGCCCTGGTTGGAAC 1422
QY 224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAsp 243
Db 1423 AGTGGCTGTGGGAAGACACACAGTCCAGCTGATGCAGAGGCTCTATGACCCACAGAC 1482
QY 244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
Db 1483 GGGATGCTCAGTGTGTGAGCAGCATATTAGGACCAATAATGTAAAGTTTCTACGGGAA 1542
QY 264 HistGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283
Db 1543 ATCAATGGTGTGGTGTGAGTACAGAACCTGTATGTTGCCACCATGATGCTGAAACAT 1602
QY 284 LysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303
Db 1603 CGCTATGCCCTGAAATGTCCACATGGATGAGATTGAGAAGCTGCAAGGAAGCCAT 1662
QY 304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyLysGly 323
Db 1663 GCCTATCACTTTATCATGAACCTCCCTCATAAATTTGACACCCCTGGTGTGGAGAGAGGG 1722
QY 324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343
Db 1723 GCCCAGTGTAGTGTGGCAGAGCAGAGATGCCCATTCAGTGCCTTGGTGTGGCAAC 1782
QY 344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
Db 1783 CCCAAGATCTCTGCTGGATGAGGCCACGTCAGCTTGGACACAGAAAGGAGCAGTG 1842
QY 364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValAlaHisArg 383
Db 1843 GTTCAGGTGGCTCTGGATAAGGCCAGAAAAAGGTCCGACCACCATTTGTATGATGCTCATCGT 1902
QY 384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403
Db 1903 TTGTACAGTGTGTAATGCTGATGCTCATCGTGGTTCGATGATGAGTATGTCGAG 1962
QY 404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423

Db 1963 AAGAGAAATCATGATCACTCATGAAGAGAAAGGCAATTTACTTCAAACTTGCACAATG 2022
QY 424 Gln-----AspIleLysLysAlaAspGluGlnMetGluSerMetThr 437
Db 2023 CACACAGCAGGAATAAGTGTGAATAGAAAATGACGTGATGAATCAAAAGTGAAT 2082
QY 438 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal----- 452
Db 2083 GATGCCCTTGGAAATGCTTCAATGATTCAGATCCAGTCTAATAAGAAAAAGATCAACT 2142
QY 453 ---LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys 470
Db 2143 CGTAGGAGTGTCCGTGGATCACAGGCCAACAGACAGAAAGCTTAGTACCAAGAGGCTCTG 2202
QY 471 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuLysProGluTyr 490
Db 2203 GATGAAGTATACCTCCAGTTCTTTTGGAGGATTTATGAAGCTAAATTTAACTGAATGG 2262
QY 491 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 510
Db 2263 CCTATTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2322
QY 511 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 529
Db 2323 GCAATAATATTTTCAAGAGATTTAGGGGTTTTTACAGAATTCATGATCTGTAACAAAA 2382
QY 530 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 549
Db 2383 CGACAGATAGTAGTAACTTGTTCCTACTATTGTTTCTAGCCCTTGGAAATTTATCTTTAT 2442
QY 550 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 569
Db 2443 ACATTTTCTCTCAGGTTTCACATTTGGCAAGCTGGAGAGATCTCCACAGCGGCTC 2502
QY 570 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 589
Db 2503 CGATACATGTTTTCGATCCATGCTCAGACAGGATGTGAGTTGTTGATGACCCCTAAA 2562
QY 590 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyVala 609
Db 2563 AACACACATGGAGCATTTGACTACAGGCTGCCAATGTAGTCTGCTCAAGTTAAAGGGCT 2622
QY 610 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629
Db 2623 ATAGTTTCCAGGCTGCTGTAAATCCAGAAATATAGCAATCTTGGACAGGAAATATT 2682
QY 630 IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 649
Db 2683 ATATCCTTCATCTATGTTGGCAACTAACACTGTTACTTCTTAGCAATTTGACCATCAT 2742
QY 650 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 669
Db 2743 GCAATGACGAGGAGTGTGAAATGAAATGTTGCTGCGAACAGCACTGAAAGATAGAAA 2802
QY 670 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 689
Db 2803 GAATGAGAGTCTGGGAGATCGCTACTGAAGCAATAGAAAACTTCCGAAACGTTGTT 2862
QY 690 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 709
Db 2863 TCTTTGACTCAGGACGAGGATTTGAACATATGTATGTCTCAGAGTTTGCAGGTACCATAC 2922
QY 710 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 729
Db 2923 AGAACTCTTTGAGAAAGCACACATCTTTTGGAAATTTACATTTCTTCCCTTCCAGGCAATG 2982
QY 730 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyValatTyrLeuIleGlnAlaGlyArg 749
Db 2983 ATGATTTTCTCTATGCTGATGTTTCCGTTTGGAGCTTACTTGTGGGCACATAAATC 3042
QY 750 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 769


```

Db 3043 ATGAGCTTTGAGGATGTTCTGTTAGTATTTTTCAGCTGTTGTTCTTTGGTGCCATGGCCGCTG 3102
QY 770 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 789
Db 3103 GGGCAAGTCAGTTCAITTTGCTCTGCTACTGTCCTCAAGCCCAAAATATCAGCAGCCACATC 3162
QY 790 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 809
Db 3163 ATCATGATCATTTGAAAACCCCTTTGATTGACAGCTACAGCAGCGAAGCCCTAAATGCCG 3222
QY 810 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 829
Db 3223 AACACATTGGAAGAATGTGCATTTGGTGAAGTGTATTCAACTATCCACCCGACCG 3282
QY 830 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 849
Db 3283 GACATCCAGTGTCTCAGGAGCTGAGCTGGAGTGAAGAAGGCCAGACGCTGGCTCTG 3342
QY 850 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 869
Db 3343 GTGGGAGCAGTGGCTGTGGGAAGAGCACAGTGTCTCCAGCTCTCGAGCGGTTCTACGAC 3402
QY 870 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTyr 889
Db 3403 CCTTGGCAGGGAAGTGTCTGTGATGGCAAGAATAAAGCAGCTGAATGTTCACTGG 3462
QY 890 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909
Db 3463 CTCGAGACACCTGGGCATCGTGTCCAGGAGCCCATCTCTGTTGACTGCACATTTGCT 3522
QY 910 GluAsnIleAlaTyrGlyAspAsnSerArgValProLeuAspGluLysLysGluAla 929
Db 3523 GAGAACATTTGCTATGAGACACACAGCCGGTGTGTACAGGAGAGATTGTGAGGGCA 3582
QY 930 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 949
Db 3583 GCNAGAGGCCAACATACATGCTCTCATCGAGTCACTGCTTAATAATATACACTAA 3642
QY 950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 969
Db 3643 GTAGGAGACAAAGAACTCAGCTCTCTGTGGTGGCCAAACACGCAATGGCATGCTCGT 3702
QY 970 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989
Db 3703 GCCTTTGTTAGACAGCTCATATTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACA 3762
QY 990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009
Db 3763 GAAAGTGAAAGGTTGTCCAAAGAGCCCTGGCAAAAGCCAGAGGCGCACCTGCAT 3822
QY 1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1029
Db 3823 GTGATGCTCACCCTGTCCACATCCAGATCCAGATTAATAGTGTGTTTCAGAT 3882
QY 1030 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1049
Db 3883 GGCAGAGTCAAGGAGCATGCGCAGCATCAGCAGCTGTGGCACAGAAAGCATCTATTT 3942
QY 1050 LysLeuValAsnAlaGlnSer 1056
Db 3943 TCAATGCTCAGTGTCCAGGCT 3963

```

RESULT 3

US-08-181-471-2

; Sequence 2, Application US/08181471

; Patent No. 5641508

; GENERAL INFORMATION:

; APPLICANT: Li, Lingna

; APPLICANT: Lishko, Valeryi K.

; TITLE OF INVENTION: METHOD FOR DELIVERING BENEFICIAL

; TITLE OF INVENTION: COMPOSITIONS TO HAIR FOLLICLES

; NUMBER OF SEQUENCES: 3

; CORRESPONDENCE ADDRESS:

;

```

; ADDRESSEE: Thomas Fitting
; STREET: 12526 High Bluff Drive, Suite 300
; CITY: San Diego
; STATE: CA
; COUNTRY: USA
; ZIP: 92130
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/181,471
; FILING DATE: 13-JAN-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/041,553
; FILING DATE: 02-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: ANT0029P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-792-3680
; TELEFAX: 619-792-8477
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4646 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 425..4267
; US-08-181-471-2

```

Alignment Scores:

```

Pred. No.: 0 Length: 4646
Score: 3100.50 Matches: 595
Percent Similarity: 75.54% Conservative: 211
Best Local Similarity: 55.76% Mismatches: 218
Query Match: 58.17% Indels: 43
DB: 1 Gaps: 6

```

US-09-873-409-4 (1-1058) x US-08-181-471-2 (1-4646)

```

QY 4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluVal 23
Db 1136 TCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCTGGAGCAGTAGCTGAAGAGTC 1195
QY 24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43
Db 1196 TTGGCAGCAATTAGAAGCTGCTGATTGCAATTTGGAGGCAAAAGAAAGAACTTGA 1249
QY 44 PheLeuLeuAsnIleThrArgTyrAlaTyrPheTyrPheProGlnTrpLeuLeuSerCys 63
Db 1249 ----- 1249
QY 64 ValLeu**PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
Db 1250 -----AGGTACACAAATAATTAGAGAGAGCTAAAGAAATTGGGATTAAG 1294
QY 84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
Db 1295 AAAGCTATTACAGCAATATTCTATAGGTGCTGCTCTCTGCTGATCTATGATCTTAT 1354
QY 104 GlyLeuAlaPheThrTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123
Db 1355 GCTCTGGCCTTCTGGTATGGGACCACCTTGGTCTCTCAGGGAA-----TATTCTATT 1408

```



```

Db 3569 GACATCCAGTCTTCAGGAGCTAGCCTGGAGGTGAAGAGGCCAGAGCGCTGGCTCTG 3628
Qy 850 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 869
Db 3629 GTGGGAGCAGTGGCTGTGGGAAGACACAGTGGTCCAGCTCTCTGGAGCGTTCTACGAC 3688
Qy 870 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 889
Db 3689 CCCTTGGCAGGGAAGTGGCTGTGATGGCAAGAAATAAAGGACCTGAATGTTCAGTGG 3748
Qy 890 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909
Db 3749 CTCGAGACACCTGGCATCGTGTCCAGAGCCCATCTCTGTTGACGACCATGCT 3808
Qy 910 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 929
Db 3809 GAGAACATTGCTATGGAGACAAACAGCCGGTGGTGTGCACAGGAAGAGATCGTGAGGGCA 3868
Qy 930 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 949
Db 3869 GCAAAGGAGGCCAACATACATGCTTTCATCGAGTCACTGCCTAATAATATAGCACTAA 3928
Qy 950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 969
Db 3929 GTAGGAGACAAAGAACTAGCTCTCTGGTGGCCAGAAACAGCATTCGCTAGCTCGT 3988
Qy 970 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989
Db 3989 GCCTTGTAGACAGCTCATATTTGCTTTTGATGAAGCCAGCTCAGCTCTGGATACA 4048
Qy 990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009
Db 4049 GAAAGTGAAAGGTTGTCCAAGAGCCCTGGCAAAAGCCAGAGAGCCGACCTGCATT 4108
Qy 1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1029
Db 4109 GTGATGCTCACCGCTGTCCACCATCCAGATTCAGACTTAATAGTGGTGTTCAGAAAT 4168
Qy 1030 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1049
Db 4169 GGCAGATCAGAGAGATGGCAGCATCAGCAGCTGCTGCGCACAGAAAGCATCTATTTT 4228
Qy 1050 LysLeuValAlaGlnInser 1056
Db 4229 TCAATGTCAGTGTCCAGGCT 4249

RESULT 4
5206352-3
; Patent No. 5206352
; APPLICANT: Roninson, Igor B.; Pastan Ira H.; Gottesman,
; Michael M.
; TITLE OF INVENTION: COMPOSITIONS FOR CLONES CONTAINING DNA
; SEQUENCES ASSOCIATED WITH MULTIDRUG RESISTANCE IN HUMAN CELLS
; NUMBER OF SEQUENCES: 4
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/622,836
; FILING DATE: 24-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 892,575
; FILING DATE: 01-AUG-1986
; APPLICATION NUMBER: 845,610
; FILING DATE: 28-MAR-1986
; SEQ ID NO:3
; LENGTH: 4669
5206352-3

Alignment Scores:
Pred. No.: 0 Length: 4669
Score: 3100.50 Matches: 595
Percent Similarity: 75.54% Conservative: 211
Best Local Similarity: 55.76% Mismatches: 218
Query Match: 58.17% Indels: 43
DB: 6 Gaps: 6

```

```

US-09-873-409-4 (1-1058) x 5206352-3 (1-4669)
Qy 4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValaGluGluVal 23
Db 1136 TCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCTGGAGCAGTAGCTGAAGAGGTC 1195
Qy 24 LeuSerSerIleArgThrValIleAlaPheAArgAlaGlnGluLysGluLeuGlnAArgSer 43
Db 1196 TTGGCAGCAATTAGAACTGTGATTGCTTTGGAGGACAAAAGAAAGAACTTGAA----- 1249
Qy 44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63
Db 1249 ----- 1249
Qy 64 ValLeu**PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
Db 1250 -----AGGTACAACAATAATTTAGAAGAGCTAAAGAAATTTGGGATAG 1294
Qy 84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
Db 1295 AAAGCTATTACAGCCAATATTCTATAGGTGCTCTTCTGCTGATCTATGCTAT 1354
Qy 104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123
Db 1355 GCTCTGGCTTCTGSPATGGGACCCTTGGTCTCTCAGGGGAA-----TATTCTATT 1408
Qy 124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
Db 1409 GGCAGTACTCTACTGTATTCTTCTGTATTAAATTTGGGGCTTTTAGTTGGACAGGCA 1468
Qy 144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163
Db 1469 TCTCCAAGCATTTGCAAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1528
Qy 164 IleAspLysIleProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
Db 1529 ATTGATAATAAGCAAGTATTGACAGTATTGCAAGAGTGGGCAAAACCAAGATAATATT 1588
Qy 184 GlyGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
Db 1589 AGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTATCCCATCTCGAAAGAGTTAAG 1648
Qy 204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
Db 1649 ATCTTGAAGGCCCTGAACCTGAAGGTGCAGAGTGGCAGAGCGTGGCCCTGGTTGAAAC 1708
Qy 224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243
Db 1709 AGTGGCTGTGGGAAGAGACAAACAGTCCAGCTGATCGAGAGGCTCTATGACCCCAAG 1768
Qy 244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
Db 1769 GGATGCTCAGTGTGATGACAGAGATATTAGACCATTAATGTAAGTTCACGGGAA 1828
Qy 264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283
Db 1829 ATCATTGGTGTGGTGTGAGTCAGGAACCTGTATTGTTGGCCACCAGTATGCTGAAACATT 1888
Qy 284 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303
Db 1889 CGCTATGGCGGTGAAATGTCACCATGTGATGATTTGAGAAAGCTGTCAAGGAAGCCAAT 1948
Qy 304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323
Db 1949 GCCTATGACTTTATCATGAACCTGCTCAATAATTTGACACCCCTGTTGGAGAGAGGG 2008
Qy 324 AlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343
Db 2009 GCCCAGTTGAGTGGTGGGAGAGAGAGATGCCATTGACGTCGCCCTGGTTGCGAAC 2068
Qy 344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363

```


Db 2069 CCCAAGATCTCTCTGGATGAGGCCACCTCAGCCCTTGGACACAGAAAACCGAAGCAGTG 2128
Qy 364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyValGlyThrThrIleValValAlaHisArg 383
Db 2129 GTTCAGGTGCTCTGGATAGGCCAGAAAAGTTCGGACCACTTGTGATGCTCATCGT 2188
Qy 384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403
Db 2189 TTGTCTACAGTTCGTAATGCTGAGCTCATCGCTGCTTCGATGATGAGTCAATTGTGGAG 2248
Qy 404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423
Db 2249 AAAAGGAATCATGATGAACCTCATGAAGAGAAAGGCCATTTACTTCAAACTTGTCACAATG 2308
Qy 424 Gln-----AspIleLysLysAlaAspGluGlnMetGluSerMetThr 437
Db 2309 CAGACACAGAAATGAAGTTGAATTAAGAAATGCACGCTGATGATCCAAAGTGAAT 2368
Qy 438 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal----- 452
Db 2369 GATGCCCTTGGAAATGCTTCAAATGATTAAGATCCAGTCTAATAAGAAAAGATCAACT 2428
Qy 453 ---LysSerIleLys---SerAspPheIleAspLysAlaGluSerThrGlnSerLys 470
Db 2429 CGTAGGAGTGTCCGTGATCACAAGCCCAAGCAAGAGCTTAGTACCAAGAGGCTCTG 2488
Qy 471 GluIleSerLeuProGluValSerLeuLysIleLeuLysLeuAsnLysProGluTyr 490
Db 2489 GATGAAAGTATACCTCCAGTTCTCTTTGGAGGATTAAGCTTAAATTTAACTGAATGG 2548
Qy 491 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 510
Db 2549 CCATATTTTGTGTGTGATTTTGTGCCATTAATAATGAGGCCCTGCAACGACATTT 2608
Qy 511 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AenAspLysThrThrLeu 529
Db 2609 GCAATAATATTTTCAAAGATATATAGGGGTTTTTACAAGAAATGATGATCTCTGAAACAAA 2668
Qy 530 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 549
Db 2669 CGACAGATAGTAACTTGTCTTACTATTTGTTCTAGCCCTTGGAAATATTTCTTTTATT 2728
Qy 550 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 569
Db 2729 ACATTTTCTTCAGGTTTCACATTTTGGCAAGCTGGAGAGATCTCCACCAAGCGCTC 2788
Qy 570 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaThrPheAspGluLysGlu 589
Db 2789 CGATACATGTTTCCGATCCATCTCAGACAGGATGTGAGTTGGTTGATGACCCCTAAA 2848
Qy 590 AsnSerThrGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 609
Db 2849 AACACCACTGGAGCATGACTTACCAGGCTCGCCAAATGATGCTCAAGTTAAAGGGGCT 2908
Qy 610 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629
Db 2909 ATAGTTTCCAGGCTTGTGTAATTTACCAGATATAGCAATCTTGGACAGGAATATT 2968
Qy 630 IleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeu 649
Db 2969 ATATCTTCTATCTATGTTGGCAACTAACACTGTACTCTTAGCAATTTGATCCCATCATT 3028
Qy 650 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 669
Db 3029 GCAATAGCAGGAGTTGTGAATGAAATGTTGTGGCAAGCACTGAAGAGATAAGAAA 3088
Qy 670 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 689
Db 3089 GAACTAGNAGTGTGGGAGATCGCTACTGTAAGCAATAGAAAACCTCCGAAACCGTTGTT 3148
Qy 690 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 709
Db 3149 TCTTTGACTCAGGAGCAGAGAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATC 3208

Qy 710 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 729
Db 3209 AGAAATCTTTGAGGAAAGCACACATCTTTGGAAATTTACATTTTCTTCCACCGGCAATG 3268
Qy 730 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyValaTyrIleuIleGlnAlaGlyArg 749
Db 3269 ATGTATTTTCTTATCTCGATGTTTCCGGTTTGGAGCCTTACTTGTGGCACATAAACTC 3328
Qy 750 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 769
Db 3329 ATGAGCTTTGAGATGTTCTGTAGTATTTTTCAGCTGTGTCTTTGGTGCATGGCCGTG 3388
Qy 770 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaHisLeu 789
Db 3389 GGGCAAGTCAGTTTCATTTGCTCTGACTATGCCAAAGCCAAATATCAGCAGCCCATC 3448
Qy 790 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 809
Db 3449 ATCATGATCATTTGAAAAAACCCCTTTGATTGACAGCTACACGACGGAAGCCCTAATGCCG 3508
Qy 810 AspThrCysGluGluGlnLeuPheArgGluValSerPhePheTyrProCysArgPro 829
Db 3509 AACACATTTGAAGAAATGTCACATTTGTGTAAGTTGTTCAACTATCCACCCGACCG 3568
Qy 830 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 849
Db 3569 GACATCCAGTGTCTCAGGACTGAGCCTGGAGGTGAAGAAGGCCAGACGCTGGCTCTG 3628
Qy 850 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 869
Db 3629 GTGGGACGAGTGGCTGTGGAAAGACACAGTGTGCTCAGCTCTCTGGAGCGGTTCTACGAC 3688
Qy 870 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTyr 889
Db 3689 CCCTTGGCAGGAAGTGTCTGTGATGGCAAGAAATAAAGCGACTGAATGTTCACTGG 3748
Qy 890 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909
Db 3749 CTCGAGCACACCTGGGCATCGTGTCCAGAGGCCCATCTCTGTTGATGTCAGCATTTGCT 3808
Qy 910 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 929
Db 3809 GAGAAATTTCCCTATGAGACAACAGCCGGTGTGTGTACAGGAAGAGATCGTGAGGGCA 3868
Qy 930 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 949
Db 3869 GCAAGAGGCCCAACATACATGCTTCTCATCGATGCTACTGCTTAATAATATAGCCTAAA 3928
Qy 950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 969
Db 3929 GTAGGACAAAGAACTCAGCTCTCTGTGGTGGCCAGAAACACGCGATGTCATGCTCGT 3988
Qy 970 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989
Db 3989 GCCCTTTAGACAGCCTCATATTTTGTCTTTTGGATGAAGCCAGTCCAGCTCTGGATACA 4048
Qy 990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009
Db 4049 GAAAGTGAAGAGTTGTCCAAAGCCCTGGCAAGCCAGAGAGCCGACCTGTCATT 4108
Qy 1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1029
Db 4109 GTGATTTCTCAGCCCTGTCCACCATCCAGATGCAGACTTAATAGTGTGTTTCAAGAT 4168
Qy 1030 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1049
Db 4169 GGCAGAGTCAAGAGAGTGGCAGCATCAGCAGCTGTGGCAGAGAAAGGATCTATTTT 4228
Qy 1050 LysLeuValAsnAlaGlnSer 1056
Db 4229 TCAATGCTAGTGTCCAGGCT 4249

STATE: D.C.
 COUNTRY: U.S.A.
 ZIP: 20005-5701
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/793,610
 FILING DATE: 07-MAR-1997
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: DE P 44 31 973.8
 FILING DATE: 08-SEP-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: DE 195 03 952.1
 FILING DATE: 07-FEB-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/EP95/03175
 FILING DATE: 10-AUG-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Berman, Richard J.
 REGISTRATION NUMBER: 39,105
 REFERENCE/DOCKET NUMBER: P1614-7007
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202)638-5000
 TELEFAX: (202)638-4810
 INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:
 LENGTH: 9318 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: circular

MOLECULE TYPE: DNA

US-08-793-610-6

Alignment Scores:

Pred. No.:	Score:	Length:	Matches:
0	3100.50	9318	595
Percent Similarity:	75.54%	Conservative:	211
Best Local Similarity:	55.76%	Mismatches:	218
Query Match:	58.17%	Indels:	43
DB:	2	Gaps:	6

US-09-873-409-4 (1-1058) x US-08-793-610-6 (1-9318)

Qy	4	SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluVal	23
Db	2487	TCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCTGGAGCAGTAGCTGAAGAGGTC	2546
Qy	24	LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer	43
Db	2547	TTGGCAGCAATTAGAACTGTGATTTGGAGGACAAAGAAAGAACTTGAA-----	2600
Qy	44	PheLeuLeuAsnIleThrArgTyrAlaTyrPheProGlnTrpLeuLeuSerCys	63
Db	2600	-----	2600
Qy	64	ValLeu***PheValArgTyrThrGlnAsnLeuLysAspPheGlyIleLys	83
Db	2601	-----AGGTCAACAAAAATTTAGAAAGCTAAAGAAATTTGGATAAAG	2645
Qy	84	ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPheMetAsnGlyThrTyr	103
Db	2646	AAAGCTATTACAGCCCAATATTTCTATAGGTGCTCTTCTCGTGTATGCTATGCTAT	2705
Qy	104	GlyLeuAlaPheTrpTyrGlyThrSerLeuIleAsnGlyGluProGlyTyrThrIle	123
Db	2706	GCTCTGGCCCTTCTGGTATGGACACCTTGGTCTCTCAGGGAA-----TATCTATT	2759
Qy	124	GlyThrValLeuAlaValPheSerValIleHisSerTyrCysIleGlyAlaAla	143
Db	2760	GGACAAGTACTCACTGTATTCTTTCTGTATTAAATTGGGGCTTTTAGTGTGGACAGCA	2819

Qy	144	ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal	163
Db	2820	TCTCCAAGCATTGAAGCATTTCGAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA	2879
Qy	164	IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle	183
Db	2880	ATTGATAATAAGCCAAAGTATTACAGCTATTGAAAGGTGGGACCAAAACAGATAATAT	2939
Qy	184	GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys	203
Db	2940	AAGGGAAATTTGGAATTCAGAAATGTTCACTTCACTTACCCATCTCGAAAGAGTTAG	2999
Qy	204	IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu	223
Db	3000	ATCTTGAAGGGCTGAACCTGAAGGTGCAGAGTGGGCAGACGCTGCTGGTGGTGAAC	3059
Qy	224	AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp	243
Db	3060	AGTGCTGTGGGAAGAGCACACAGTCCAGCTGATGAGAGGCTCTATACCCACAGAG	3119
Qy	244	GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp	263
Db	3120	GGAGTGTGAGTGTGATGCGACAGGATATTAGCCATTAATGTAAAGTTCTACGGGAA	3179
Qy	264	HisIleGlyValValSerGlnProValLeuPheGlyThrThrIleSerAsnAnile	283
Db	3180	ATCATTTGGTGTGGTGTGAGTCAGGAACCTGTATTGTTGCCACCCAGATAGCTGAAACAT	3239
Qy	284	LysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn	303
Db	3240	CGCTATGGCGGTGAAATGTCCATGATGATGAGAAAGCTGTCAAGGAAGCAAT	3299
Qy	304	AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly	323
Db	3300	GCCTATGACTTTATCATGAACCTGCTCATTAATTTGACACCTGTTGAGAGAGAGGG	3359
Qy	324	AlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn	343
Db	3360	GCCAGTTGAGTGGTGGCAGAAAGCAGAGAGTGCCTATGTCACGTCCCTGTTGCAAC	3419
Qy	344	ProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla	363
Db	3420	CCCAAGATCTCTGCTGATGAGGCCAGCTCAGCTGAGCCCTGGACACAGAAAGCAGTG	3479
Qy	364	ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg	383
Db	3480	GTTGCTGCTGCTGATGAGCCAGAAAGGTGCGACCACTTGTGATGCTCATCGT	3539
Qy	384	LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu	403
Db	3540	TTGCTTACAGTTGCTAATGCTGACGTCATCGCTGCTGCTGATGATGAGTGTGGAG	3599
Qy	404	LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer	423
Db	3600	AAAGGAAATCATGATGAATCATGAAAGAGAGAGGCACTTACTTCAAACTGTCAAAATG	3659
Qy	424	Gln-----AspIleLysLysAlaAspGluGlnMetGluSerMetThr	437
Db	3660	CACACAGCAGGAAATGAAGTTGAATTAGAAATGTCAGCTGATGAATCCAAAGAGTGAAT	3719
Qy	438	TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal-----	452
Db	3720	GATGCTTGGAAATGCTTCAATGATTCAGATCCAGTCTATATAGAAAAAGATCAACT	3779
Qy	453	---LysSerIleLys---SerAspPheIleAspLysAlaGluSerThrGlnSerLys	470
Db	3780	CGTAGGAGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	3839
Qy	471	GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuLysProGluTrp	490
Db	3840	GATGAAAGTATACCTCCAGTTCTTCTTTGGAGGATTAAGAACTTAAATTAAGTGAATG	3899

QY 491 ProPheValValLeuGlyThrLeuAlaSerValLeuAenGlyThrValHisProValPhe 510
Db 3900 CTTATTTTGTGTGTGTATTTTGTGCTTAAATTAATGAGGCGCTGCAACACGCAATTT 3959
QY 511 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AenAspLysThrLeu 529
Db 3960 GCAATAATATTTTCAAGATATATAGGGGTTTTTACAGAAATGATGATCTGAAACAAAA 4019
QY 530 LysHisAspAlaGluIleIleThrSerMetIlePheValIleLeuGlyValIleCysPheVal 549
Db 4020 CGACAGAAATAGTAATGTTTCTACTATTGTTCTAGCCCTTGGAAATATTCTTTTATT 4079
QY 550 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 569
Db 4080 ACATTTTCTCTCAGGGTTTTCACATTTTGGCAAGCTGAGAGATCTCTCAACCAAGCGGCTC 4139
QY 570 ArgHisIleAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGluLysGlu 589
Db 4140 CGATACATGTTTTCGATCCATCTCAGACAGGATGTGAGTTGGTTGATGACCCATAA 4199
QY 590 AsnSerThrGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyVala 609
Db 4200 AACACCTGGAGCATTTGACTACAGGCTGCCAATGATGCTGCTCAAGTTAAGGGGCT 4259
QY 610 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629
Db 4260 ATAGTTTCCAGGCTTGCTGTAATTAACCAAGATATAGCAAAATCTTGGGACAGGAATATT 4319
QY 630 IleSerPheIleTyrGlyTyrGluMetThrPheIleIleLeuSerIleAlaProValLeu 649
Db 4320 ATATCTTCTATGTTGGCACTAACACTGTACTTCTTAGCAATGTACCCATCAT 4379
QY 650 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 669
Db 4380 GCAATACAGGAGTGTGTGAATGAAATGTGTCTGGACAGCACTGAAGATAGAAA 4439
QY 670 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 689
Db 4440 GAACTAGAAGTGTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTGTT 4499
QY 690 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 709
Db 4500 TCITTGACTCAGGAGCAGAGATTGGAACATGATGATGCTCAGAGTTTGCAGGTACCATAC 4559
QY 710 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 729
Db 4560 AGAACTCTTGAGAAAGCACACATCTTGGAAATACATTTCTTCCACCGCAATG 4619
QY 730 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 749
Db 4620 ATGTATTTTCTCTATGCTGGATGTTCCGGTTTGGAGCTACTTGTGGCACATAAACTC 4679
QY 750 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 769
Db 4680 ATGAGCTTTGAGGATGTTCTGTAGTATTTTACGCTGTTCTTGTGTCATGGCGGTG 4739
QY 770 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 789
Db 4740 GGGCAAGTCAGTTCATTTGCTCTGATATGCCAAACCAATATATCAGAGCCCATC 4799
QY 790 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 809
Db 4800 ATCATGATCATTTGAAAAAACCCTTTTCATTGACAGCTACAGCAGGAAGGCTTAATCGC 4859
QY 810 AspThrCysGluGlyAnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 829
Db 4860 AACCATTTGGAAGGAATGTACATTTTGGTGAAGTGTATTTCAACTATCCACCCCGACCG 4919
QY 830 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 849
Db 4920 GACATCCAGTGTCTTACGAGACTGAGCTGGAGGTGAAGAGGCCAGAGCTGGCTCTG 4979
QY 850 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 869

Db 4980 GTGGGACGAGTGTGTGGAAAGACACAGTGTGTCCAGCTCTCTGGAGCGGTTCTACGAC 5039
QY 870 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAnValGlnTyr 889
Db 5040 CCTTGGCAGGAAGTGTGTGTATGGCAAGAATAAAGCAGCTGAATGTTTCAGTGG 5099
QY 890 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909
Db 5100 CTCGAGACACACTGGGCATCGTGTCCAGAGGCCCATCTCTGTTGACTGCAGCATTTGCT 5159
QY 910 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 929
Db 5160 GAGAACAATTTGCTTATGGAGACAAACAGCCGGTGTGTACAGGAAGAGATCGTGAGGSCA 5219
QY 930 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 949
Db 5220 GCAAGAGGCCCAACATACATGCTTCTATCGAGTCTACTGCTTAATAATATAGCACTAAA 5279
QY 950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 969
Db 5280 GTAGGAGACAAAGAACTCAGCTCTCTGTGGGCCAGAAACACGCAATGCCATAGCTCGT 5339
QY 970 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989
Db 5340 GCCTTGTGTAGACAGCTCATATTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACA 5399
QY 990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009
Db 5400 GAAAGTGAAGAGTTGTCAAGAAGCCCTGGACAAAGCCAGAGAGCCGCACTGCAAT 5459
QY 1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1029
Db 5460 GTGATTTGCTCACCGCTCTCCACCATCCAGAAATGCAGACTTATAGTGGTGTTCAGAAT 5519
QY 1030 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1049
Db 5520 GGCAGAGTCAAGAGCATGCGCAGCATCAGCAGCTGTGTCACAGAAAGCAATCTATTTT 5579
QY 1050 LysLeuValAsnAlaGlnSer 1056
Db 5580 TCAATGTCAGTCCAGGCT 5600

RESULT 7
US-09-120-513-1
; Sequence 1. Application US/09120513
; Patent No. 6025160
; GENERAL INFORMATION:
; APPLICANT: Brun, Kimberly
; APPLICANT: Chenery, Richard
; APPLICANT: Ellens, Harma
; APPLICANT: Field, John
; APPLICANT: Yue, Lin
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE
; TITLE OF INVENTION: SEQUENCES ENCODING RAT MDR1B2 AND
; TITLE OF INVENTION: SCREENING METHODS THEREOF
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY:
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/120,513
; FILING DATE: 22-JUL-1998
; CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: King, William T
 REGISTRATION NUMBER: 30,954
 REFERENCE/DOCKET NUMBER: GP50008
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 610-270-5015
 TELEFAX: 610-270-5090
 TELEX:

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 4233 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: Genomic DNA

US-09-120-513-1

Alignment Scores:

Pred. No.: 0 Length: 4233
 Score: 3086.00 Matches: 601
 Percent Similarity: 74.74% Conservative: 198
 Best Local Similarity: 56.22% Mismatches: 224
 Query Match: 57.90% Indels: 46
 DB: 3 Gaps: 7

US-09-873-409-4 (1-1058) x US-09-120-513-1 (1-4233)

Qy 1 MetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAla 20
 Db 722 GTACTGACTTCTTACTATAAGAACTCCAGCTTATGCGAAGCTGGAGCAGTTGCC 781
 Qy 21 GluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeu 40
 Db 782 GAAGAAGTCTTAGCAGCCATCAGAACTGTGATTGCGTTGGAGGACAAAAGAGGAACCT 841
 Qy 41 GlnArgSerPheLeuLeuAlaIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeu 60
 Db 842 GAA----- 844
 Qy 61 LeuSerCysValLeu**PheValArgTyrThrGlnAsnLeuLysAspAlaTyrAspPhe 80
 Db 845 -----AGTACATATAAATTTAGAAAGCTAAAGAGTT 880
 Qy 81 GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPheMetAsn 100
 Db 881 GGCATTAAGAAACCCATCAGGCCAACATTTCCATAGTATTGCTACCTGTTGGTCTAT 940
 Qy 101 GlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAlaGlnGluProGly 120
 Db 941 GCGTCTTATGCACCTGTCGATTCGGTATGGACCTCTTGGTCTCTCAATGAA----- 994
 Qy 121 TyrThrIleGlyThrValLeuAlaValPheSerValIleHisSerSerTyrCysIle 140
 Db 995 TATCTATTGGACAAGTGTCTACCGTCTCTCTCTATTTATTTGGGAGCTTCAGTATT 1054
 Qy 141 GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 160
 Db 1055 GGACATTAGCCCAACATAGAGCTTTTGCAATGCAAGAGGGGCGAGCTATGAATC 1114
 Qy 161 PheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysPro 180
 Db 1115 TTCAAGATAATTGATAATGAGCCCAAGCAATTCAGACGTTCTCTCAACCAAGGGGACACAACCA 1174
 Qy 181 GluSerIleGluThrValGluPheLysAsnValSerPheAsnTyrProSerArgPro 200
 Db 1175 GACAGTATAATGGAAATTTGGAAATTTAAATAATGTTTACTTCACTACCCATCACGAAGT 1234
 Qy 201 SerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu 220
 Db 1235 GAAGTTAAGATCTTGAAGGCCCTCAACCTGAAGTGAAGAGCGGAGACGGTAGCCCTG 1294
 Qy 221 ValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAsp 240
 Db 2369 ACCAAGCGACTCCGATACATGTTCTTCAAAATCCATCTCTCGCAGCAGGATATAAGCTGGTTT 2428

Db 1295 GTTGGCAACAGTGGCTGTGGGAAAAGCACAACTGTCCAGCTGCTGCAGAGGCTCTACGAC 1354
 Qy 241 ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis 260
 Db 1355 CCCATAGAGGGCGAGGTAGTATCGACGACAGGACATCAGGACCATCAATGTGAGGTAT 1414
 Qy 261 TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSer 280
 Db 1415 CTGCGGAAATCATTTGGGTGGTGAAGTCAAGAACCCGCTGCTGTTGCCACCATGTC 1474
 Qy 281 AsnAsnIleLysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArg 300
 Db 1475 GAACAATTCGCTATGGCCGAGAAACGTCACATGATGAGATAGAGAAAGCTGCAAG 1534
 Qy 301 GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly 320
 Db 1535 GAAGCCAATGCTATGACTTCATGAACTGCCCCACAAATTTAACACCTGTTGGT 1594
 Qy 321 GluLysGlyAlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu 340
 Db 1595 GAGAGAGGGCGCAGCTGAGTGGGACAGAAACAGAGATCGCCATGCCGGGCCCTG 1654
 Qy 341 ValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 360
 Db 1655 GTCCGCAACCCCAAGATCCTTTTGGATGAGCCACGTCAGCCTTGACACAGAAAGC 1714
 Qy 361 LysSerAlaValGlnAlaLeuGluLysAlaSerLysGlyArgThrThrIleValVal 380
 Db 1715 GAAGCGTGTTCAGGCCCTCTGGATAAGGCTAGAGAGGCGGAGCCACCATTTGTGATA 1774
 Qy 381 AlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMet 400
 Db 1775 GCTCACCGTGTGTACAGTGGCAATGCTGACGTCAATTTGCTGGTTTGTGTTGGTGTGTC 1834
 Qy 401 LeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeu 420
 Db 1835 ATTTGGAGCAAGGAATCATGAAGAGCTCATGAAGAGAGGGCAATTTACTTCAAACT 1894
 Qy 421 ValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThr 440
 Db 1895 GTCATGACACAG---ACTAGAGGAATGAATTTGAACAGGAAATTAATGCTTATGAATCC 1951
 Qy 441 GluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle 460
 Db 1952 CAAAGTGACACTGGTGCCTCTGAGTTGACTTCAGAAAAATCA---AAATCTCCTTTAATA 2008
 Qy 461 AspLysAla-----GluGluSerThrGlnSer 469
 Db 2009 AGGAGATCAATTCGCAGAAAGTATCCAGAAAGACAAAGACAGAGAGAGACTTAGTTCG 2068
 Qy 470 LysGlu-----IleSerLeuProGluValSerLeuLysIleLeuLysLeuAsn 486
 Db 2069 AAAGAGATGTGGATCAAGATGTGCTATGTTCTCTTTGGCAGATCTTAAGCTAAT 2128
 Qy 487 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 506
 Db 2129 ATTAGTAATGGCCCTATTATTAGTTGGGTGTACTTTGTCTGTATAAATGGGTGCATA 2188
 Qy 507 HisProValPheSerIlePheAlaLysIleIleThrMetPheGly---AsnAsnAsp 525
 Db 2189 CAACAGTGTGGCCATAGTGTTCAAAGATTTGAGGGGTTTTTCAAGAGACGACGAC 2248
 Qy 526 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 545
 Db 2249 CATGAACCAACCAACGGAATGTAACTTTGTTTCTCTCTCTCTCTCTCTCTCTCTCTCT 2308
 Qy 546 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 565
 Db 2309 ATTTCTTTTGTAGTACTTCTTCAAGGCTTCACATTTGGCAAGAGCTGGAGAGATCCTC 2368
 Qy 566 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 585
 Db 2369 ACCAAGCGACTCCGATACATGTTCTTCAAAATCCATCTCTCGCAGCAGGATATAAGCTGGTTT 2428

QY 786 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 805
Db 3029 GCATCCACATCATCAGCATCATTTAGAGAAATCCCGAGATTGACAGCTACAGACCGAG 3088
QY 806 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 825
Db 3089 GCGTTGAAGCCTAATGTTAGAGAGAAATGTAATTAATGAGTCACTGTTCAACTAT 3148
QY 826 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 845
Db 3149 CCCACCGACCCCAATCCAGTCTTCAGGAGCTGAGCTCGAGGTGAAGAGGGGCAA 3208
QY 846 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 865
Db 3209 ACGTTCGCTGCTGGGCGAGCAGTGGCTCGGGAAGATGACAGTGGTCCAGCTGCTCAG 3268
QY 866 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 885
Db 3269 CGCTTCTACAAACCCCATGGCTGGAAACAGTGTCTAGATGCCAAGAAATAAAACAAC 3328
QY 886 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 905
Db 3329 AACGTCAGTGGCTCCGCGCCACCTGGCATTTGTCCAGGAGCCCTCTGTTTGAC 3388
QY 906 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 925
Db 3389 TGCAGCATCCAGAGACATCGCTACGGAGACAAACAGCCGTGTCGTCTCATGAGGAG 3448
QY 926 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 945
Db 3449 ATCGTAGGGCGCGCAGGAGGCGCAACATCCACCGATTTCATCGACTCACTGCCTGAGAAA 3508
QY 946 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 965
Db 3509 TACAACACCGAGTGGGAGACAAAGGAGCTAGCTGTGGCGGGCAGAGCAGCGCATC 3568
QY 966 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuAspGluAlaThrSer 985
Db 3569 GCCATCGCGGGCGCTCGTCAGACAGCTCACATCTTACTTCTGGATGAAGCGACATCA 3628
QY 986 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 1005
Db 3629 GCTCTGATACGAGAGTGAAGAGTGGTCCAGAGCGCTGGACAAAGCCAGGAAAGGC 3688
QY 1006 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 1025
Db 3689 CGCACCTGCATTGTATCGCGCACCGCTGTCCACCATCCAGAACGCGAGACTTGTATGCTG 3748
QY 1026 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlnLeuLeuArgAsnArg 1045
Db 3749 GTGATTCAAGACGCGCAGGTCAAGGAGCAGCGCACCCACCCAGCAGCTGCTGGCCAGAAA 3808
QY 1046 AspIleTyrPheLysLeuValAsnAla 1054
Db 3809 GGCATCTATTCTCGATGGTTTCAGGCT 3835

RESULT 9

US-08-752-447-1
; Sequence 1, Application US/08752447
; Patent No. 5994088
; GENERAL INFORMATION:
; APPLICANT: Mechtner, Eugene
; APPLICANT: Roninson, Igor B
; TITLE OF INVENTION: Methods and Reagents for Preparing and
; TITLE OF INVENTION: Using Immunological Agents Specific for P-glycoprotein
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boenhen Hulbert & Berghoff Ltd.
; STREET: 300 South Wacker Drive, Seventh Floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/752,447
FILING DATE: 15-NOV-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: No. 5994088han, Kevin E
REGISTRATION NUMBER: 35,303
REFERENCE/DOCKET NUMBER: 95,1121
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-913-0001
TELEFAX: 312-913-9808
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4669 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..424
FEATURE:
NAME/KEY: CDS
LOCATION: 425..4264
FEATURE:
NAME/KEY: 3'UTR
LOCATION: 4265..4669
US-08-752-447-1
Alignment Scores:
Pred. No.: 0 Length: 4669
Score: 3073.50 Matches: 590
Percent Similarity: 74.98% Conservative: 210
Best Local Similarity: 55.30% Mismatches: 224
Query Match: 57.66% Indels: 43
DB: 2 Gaps: 6
US-09-873-409-4 (1-1058) x US-08-752-447-1 (1-4669)
QY 4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluVal 23
Db 1136 TCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCTGGAGCAGTAGCTGAAGAGTC 1195
QY 24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43
Db 1196 TTGGCAGCAATTAGAACTGTGATTGCAATTGGAGGACAAAAGAAAGAACTTGAA----- 1249
QY 44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63
Db 1249 ----- 1249
QY 64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
Db 1250 -----AGGTACAAACAAATTTAGAGAGCTAAAGAAATTGGGATAAG 1294
QY 84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
Db 1295 AAGACTATTACAGCCAATTTCTATAGGTGCTGCTTCTCTGCTGATCATCTAT 1354
QY 104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123
Db 1355 GCTGTGCGCTTCTGGTATGGACACCTTGGTCTCTCAGGGAA-----TATTCTATT 1408
QY 124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
Db 1409 GGACAAGTACTCTACTGATTCTTTCTGTATTAATTTGGGGCTTTTGTGTGACAGGCA 1468
QY 144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163

Db	1469	TC	CC	AG	CA	TT	GC	AA	AT	GC	AA	GC	AG	CA	GC	TT	AT	GA	AA	TC	TT	CA	AG	AT	1528		
Qy	164	I	e	A	s	p	L	y	s	P	r	o	S	e	r	I	e	A	s	p	A	e	n	P	e	r	183
Db	1529	A	T	T	G	A	T	A	A	G	C	A	G	T	A	T	T	G	A	G	C	T	A	T	T	1588	
Qy	184	G	l	u	G	l	y	T	h	r	V	a	L	g	h	u	P	e	L	y	s	A	s	n	V	a	203
Db	1589	A	A	G	G	A	A	T	T	G	G	A	A	T	T	G	C	A	A	T	T	G	C	A	A	1648	
Qy	204	I	e	L	e	u	L	y	S	e	r	I	e	u	A	s	n	L	e	u	A	r	g	L	e	u	223
Db	1649	A	T	C	T	T	A	A	G	G	C	T	G	A	C	T	T	A	A	G	G	T	G	C	A	1708	
Qy	224	A	s	n	G	l	y	S	e	r	T	h	r	V	a	L	g	h	u	L	e	u	L	e	u	243	
Db	1709	A	G	T	G	C	T	G	G	A	G	A	C	A	C	A	C	A	G	T	C	A	G	T	1768		
Qy	244	G	l	y	P	h	e	I	e	M	e	t	V	a	L	g	h	u	A	s	n	A	s	p	I	e	263
Db	1769	G	G	A	T	G	T	C	A	G	T	T	G	A	T	T	G	A	C	A	G	G	A	T	1828		
Qy	264	H	i	s	I	e	G	l	y	V	a	L	g	h	u	P	r	o	V	a	L	e	u	P	e	283	
Db	1829	A	T	C	A	T	T	G	T	G	T	G	A	G	T	G	A	G	A	C	T	G	T	1888			
Qy	284	L	y	S	e	r	G	l	y	A	r	g	A	s	p	V	a	L	g	h	u	M	e	t	G	303	
Db	1889	C	G	C	A	T	G	G	C	G	T	G	A	A	T	G	T	C	A	G	A	T	T	1948			
Qy	304	A	a	L	a	T	y	r	A	s	p	P	h	e	I	e	M	e	t	G	l	u	P	e	323		
Db	1949	G	C	C	A	T	G	A	C	T	T	A	T	T	G	C	A	A	T	T	G	C	A	2008			
Qy	324	A	a	L	a	G	l	m	e	t	S	e	r	G	l	y	G	l	n	L	y	s	G	l	n	343	
Db	2009	G	C	C	A	G	T	T	G	A	G	T	G	G	A	G	A	G	A	G	A	G	A	2068			
Qy	344	P	r	o	L	y	s	I	e	L	e	u	L	e	u	A	s	p	G	l	u	A	t	h	r	363	
Db	2069	C	C	A	A	G	A	T	C	T	C	T	G	A	T	G	A	G	C	C	A	G	T	2128			
Qy	364	V	a	L	g	h	a	L	a	L	e	u	G	l	u	L	y	s	A	l	a	S	e	r	383		
Db	2129	G	T	T	C	A	G	T	G	C	T	G	G	A	A	G	G	T	C	G	A	C	C	2188			
Qy	384	L	e	u	S	e	r	T	h	r	L	e	u	S	e	r	A	a	s	p	L	e	u	L	e	403	
Db	2189	T	T	T	G	T	C	A	G	T	T	G	A	T	T	G	A	T	T	G	A	T	2248				
Qy	404	L	y	S	e	r	G	l	u	L	e	u	M	e	t	A	L	a	y	s	A	r	g	L	e	423	
Db	2249	A	A	G	G	A	A	T	C	A	T	G	A	A	G	A	A	G	A	G	A	G	A	2308			
Qy	424	G	l	n	-----	-----	-----	-----																			

D	b	2549	CCTTATTTTGTGTTGGTGATATTTTGTGCCATATAAATGAGGCGCTCGAACACGACATTT	2608
Q	y	511	SerIlePheAlaLysIleThrMetPheGlyAsn--AsnAspLysThrThrLeu	529
D	b	2609	GCAATAATATTTCAAGATTATAGGGTTTTCACAGAAATGATGATCCTGAAACAAA	2668
Q	y	530	LysHisAspAlaGluLeuTyrSerMetIlePheValIleLeuGlyValIleCysPheVal	549
D	b	2669	CGACAGAAATAGTAACCTGTTTTCACATATTGTTCTAGGCCCTTGAATAATTTCTTTTATT	2728
Q	y	550	SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu	569
D	b	2729	ACATTTTCTCTCAGGTTTCACATTTGCAAGCTCGAGAGATCCTCCACAGGGCTC	2788
Q	y	570	ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTTPheAspGluLysGlu	589
D	b	2789	CGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGTTTCATGACCTAAA	2848
Q	y	590	AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla	609
D	b	2849	AACACCACTCGAGCATTTGACTACAGGCTCGCAATGATGCTGCTCAAGTTTAAAGGGCT	2908
Q	y	610	ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle	629
D	b	2909	ATAGGTTCCAGGCTGCTGTAATATCCCAAGATATAGCAAAATCTTGGGACAGGAATATT	2968
Q	y	630	IleSerPheIleTyrGlyTTPGluMetThrPheLeuIleLeuSerIleAlaProValLeu	649
D	b	2969	ATATCCTTCATCTATGTTGGCAACTAAACACTGTTTACTCTTAGCAATGTACCCATCAT	3028
Q	y	650	AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln	669
D	b	3029	GCAATAGCAGGAGTTGTTGAAATGAAAAATGTTTGTGTCGACAGCACTGAAAGATAAGAAA	3088
Q	y	670	GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal	689
D	b	3089	GAACTAGAGGTCGTGGGAAGTCGCTACTCTAGAGCAATAGAAAACTCCGAAACCGTTGTT	3148
Q	y	690	SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis	709
D	b	3149	TCCTTTGACTCAGGACGAGAAGTTGAACATATGTATGCTCAGAGTTTCAGGTACCATAC	3208
Q	y	710	ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe	729
D	b	3209	AGAAACTCTTTGAGGAAGCACACATCTTTTGAATAATACATTTTCTCCACCCAGCAATG	3268
Q	y	730	IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg	749
D	b	3269	ATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGTGGCCACATAAACTC	3328
Q	y	750	MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle	769
D	b	3329	ATGAGCTTCAGGATGTTCTGTTAGTATTTTTCAGCTGTTGTTGTTGGTGCCATGGCGTG	3388
Q	y	770	GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu	789
D	b	3389	GGGCAAGTCAGTTCAITTTGCTCTGACTATGCCAAAGCCAAAATATCAGCAGCCACATC	3448
Q	y	790	PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysAlaPro	809
D	b	3449	ATCATGATCATTTGAAAAACCCCTTTGATTTGACGTACAGCAGGAGGGCTTAATGCCG	3508
Q	y	810	AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro	829
D	b	3509	AACATCTGGNAGGAATGTCACATTTGGTGAAGTTGATTCAACTATCCACCCGACCG	3568
Q	y	830	AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe	849
D	b	3569	GACATCCAGTGTCTCAGGACTGACCTGGAGGTGAGAGGGCCAGACGCTGGCTCTG	3628
Q	y	850	ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp	869
D	b	3629	GTGGGCAGCAGTGGCTGTGGGAAGACACAGTGGTCCAGCTCTCGAGCGGTTCTACGAC	3688


```
QY 870 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 889
Db 3689 CCCTGGCAGGAAAGTGTCTGTATGTCGCAAGAAATAAGCAGCTGAATGTCAGTGG 3748
QY 890 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909
Db 3749 CTCGAGCACACCTGGGCATCGTGTCCAGAGGCCATCTCTTTGACTGCAGCATGCT 3808
QY 910 GluAsnIleAlaTrpGlyAspAsnSerArgValProLeuAspGluIleLysGluAla 929
Db 3809 GAGAACATTGCTTATGAGACAACAGCCGGGTGTGTACAGGAGAGATCGTGAGGGCA 3868
QY 930 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTrpAsnThrGln 949
Db 3869 GCAAAGGAGCCCAACATACATGCTTCATCGAGTCACGTGCTAATAATATAGCACTAAA 3928
QY 950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlnLysGlnArgLeuAlaIleAlaArg 969
Db 3929 GTAGGAGACAAAGGAACTCAGCTCTCTGTGGCCAGAAACAACGCAATGCGCATGCT 3988
QY 970 AlaLeuGlnLysProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989
Db 3989 CGCCTTGTAGACAGCCTCATATTTTGTGATGAAGCCAGCTCAGCTCTGGATACA 4048
QY 990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009
Db 4049 GAAAGTGAAAGGTTGTCCAAAGAGCCCTGGACAAACCCAGAGAGCCGCACTGCATT 4108
QY 1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1029
Db 4109 GTGATTCTCACCGCCTGTCCACCATCCAGAAATGAGACTTAATAGTGTGTTCAGAAAT 4168
QY 1030 GlyLysIleLysGluGlnGlyThrHisGlnLeuLeuAsnArgAspIleTrpPhe 1049
Db 4169 GCGAGTCAAGGAGCATGGCCCATCAGCATCAGCAGCTGTGCGACAGAAAGCACTATTTT 4228
QY 1050 LysLeuValAsnAlaGlnSer 1056
Db 4229 TCAATGTCAGTGTCCAGGCT 4249

RESULT 10
US-09-316-167-1
; Sequence 1, Application US/09316167
; Patent No. 6365357
; GENERAL INFORMATION:
; APPLICANT: Mechtner, Eugene
; APPLICANT: Roninson, Igor B
; TITLE OF INVENTION: Methods and Reagents for Preparing and
; TITLE OF INVENTION: Using Immunological Agents Specific for P-glycoprotein
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff Ltd.
; STREET: 300 South Wacker Drive, Seventh Floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/316,167
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/752,447
; FILING DATE: 15-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6365357nan, Kevin E
; REGISTRATION NUMBER: 35,303
```

```
; REFERENCE/DOCKET NUMBER: 95,1121
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-9808
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4669 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: 1..424
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 425..4264
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: 4265..4669
US-09-316-167-1

Alignment Scores:
Pred. No.: 0 Length: 4669
Score: 3073.50 Matches: 590
Percent Similarity: 74.98% Conservative: 210
Best Local Similarity: 55.30% Mismatches: 224
Query Match: 57.66% Indels: 43
DB: 6

US-09-873-409-4 (1-1058) x US-09-316-167-1 (1-4669)
QY 4 SerLeuThrSerLysGluLeuSerAlaTrpSerLysAlaGlyAlaValaGluVal 23
Db 1136 TCATTTACTGATAAGAACTCTTAGGTATGCAAAAGCTGGAGCAGTAGCTGAAGAGT 1195
QY 24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43
Db 1196 TTGGCAGCAATTAGAACTGTGATTGCTATTTGGAGGACAAAAGAAAGAACTTGAA 1249
QY 44 PheLeuLeuAsnIleThrArgTrpAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63
Db 1249 ----- 1249
QY 64 ValLeu***PheValArgTrpThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
Db 1250 -----AGGTACACAAAATTTAGAGAGAGCTAAAGAAATTTGGATAAG 1294
QY 84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
Db 1295 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTCTCTGCTGATCTATGTCATCTAT 1354
QY 104 GlyLeuAlaPheTrpTrpGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123
Db 1355 GCTGTGCGCTTCTGTTATGGGACCACCTTGCTCTCTCAGGGGAA-----TATTCTATT 1408
QY 124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
Db 1409 GGACAAGTACTACTGTATTTCTTCTGTATTAATTTGGGGCTTTTAGTTGACAGGCA 1468
QY 144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163
Db 1469 TCTCCAAGCATTTGAAGCATTTTGCAAATGCAAGAGAGGAGCAGCTTATGAAATCTTCAAGATA 1528
QY 164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
Db 1529 ATTGATAATAAGCCAAGTATTTGACGATTTCGAAGAGTGGGCACCAACAGATAATATT 1588
QY 184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
Db 1589 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAGTTAAG 1648
QY 204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValaLeuValGlyLeu 223
```


[illegible]

D	b	2729	ACATTTTTCCTTCAGGGTTTCAATTTGGCAAGAGCTGGAGAGATCCTCACCAGAGGGCTC	2718
Q	y	570	ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrrPheAspGluLysGlu	589
D	b	2789	CGATACTATGGTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTCATGACCCATAA	2848
Q	y	590	AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla	609
D	b	2849	AACACCACTGGACGATTTGACTACCAAGCTCGCAATGATGTCTCAAGTTTAAAGGGGCT	2908
Q	y	610	ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle	629
D	b	2909	ATAGTTTCCAGGCTGCTGTAAATTTACCCAGAATATAGCAAAATCTTTGGGACAGGAATAA	2968
Q	y	630	IleSerPheIleTyrGlyTrrPgluMetThrPheLeuIleLeuSerIleAlaProValLeu	649
D	b	2969	ATATCCTTCATCTATGGTTGGCAACTAAACATGTTACTCTTAGCAATTTGTACCCATCAT	3028
Q	y	650	AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln	669
D	b	3029	GCAATAGCAGGAGTTGTTGAAATGAAAAATGTTGTGTGGCAAGCACTGAAAGATAAGAAA	3088
Q	y	670	GluLeuLysHisIalaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal	689
D	b	3089	GAATAGAGGCTGCTGGAGAGTCGCTACTGAAGCAATAGAAAACTTCCGAACCCTGTGTT	3148
Q	y	690	SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis	709
D	b	3149	TCITTTGACTCAGAGCAGAAAGTTTGNACATATGTATGCTCAGAGTTTCAGGTACCATC	3208
Q	y	710	ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe	729
D	b	3209	AGAAACTCTTTTGAGGAAGCACACATCTTTTGGAAATTAATATTTCTCCACCAGCAATG	3268
Q	y	730	IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg	749
D	b	3269	ATGTAATTTTTCCTATGCTGATGTTTCCGGTTTGAGGCTACTTGGTGGCACAATAAATC	3328
Q	y	750	MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle	769
D	b	3329	ATGAGCTTTGAGGATGTTCTGTAGTATTTTTCAGCTGTGTTGTTGGTGCCATGGCCGTG	3388
Q	y	770	GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu	789
D	b	3389	GGGCAAGTCAGTTTCATTTGCTCTGACTATGCCAAGGCCAAAATATCAGCAGCCACATC	3448
Q	y	790	PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro	809
D	b	3449	ATCATATGATCATTTGAAAAAACCCTTTTGTATGTGACGCTACAGCAGGAAGGCTTAATG	3508
Q	y	810	AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro	829
D	b	3509	AACACATTTGGAAGAAATGTCAATTTGTTGAAGTTGTATTCAACTATCCACCAGCCG	3568
Q	y	830	AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe	849
D	b	3569	GACATCCAGTGCTTCAGGACCTAGCCCTGGAGGTGAAGAGGGCCAGACGCTGGCTGTG	3628
Q	y	850	ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp	869
D	b	3629	GTGGGACAGTAGTGGCTGTGGGAAGACACAGTGGTTCAGCTCTCTGAGCGGTTCTACGAC	3688
Q	y	870	ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrr	889
D	b	3689	CCCTTGGCAGGAAAGTGCTGTGTTGATGGCAAGAAATAAAGCGACTGAAATTTTCAGTGG	3748
Q	y	890	LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla	909
D	b	3749	CTCCGACACACCTGGGCATCTGTGTCCAGGAGCCCATCTCTGTTTGTAGCTGAGCATTTG	3808
Q	y	910	GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla	929
D	b	3809	GAGAAATTCCTATTCGATGACAAACAGCCGGGTGGTGTTCACAGGAGAGATCGTAGGGCA	3868


```
QY 930 AlaAsnAlaAlaAenIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 949
|||
Db 3869 GCAAGAGGCGCAACATACATGCTTCATCGAGTCACCTGCTAATAATATAGCACTAAA 3928
|||
QY 950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaAlaArg 969
|||
Db 3929 GTAGGAGCAAGGAACTCAGCTCTCTGGTGGCCAGAAACAAACGCAATTGCCATAGCTCGT 3988
|||
QY 970 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989
|||
Db 3989 GCGCTTGTTAGACAGCGCTCATATTTGCTTTGGATGAAGCCAGCTCAGCTCGATGATACA 4048
|||
QY 990 AspSerGlyLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009
|||
Db 4049 GAAAGTGAAGAGTTGTCAGAGAGCTTGGACAAAGCCAGAGAGCGCCAGCTGCAATT 4108
|||
QY 1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1029
|||
Db 4109 GTGATTGCTACCGCGCTGTCCACCATCCAGAAATCAGACTTAAATAGTGGTGTTCAGAAAT 4168
|||
QY 1030 GlyLysIleLysGlnGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1049
|||
Db 4169 GGCAGAGTCAGAGGATGCGCAGCATCAGCAGCTGCTGGCAGCAAGAGGCAATCTATTTT 4228
|||
QY 1050 LysLeuValAsnAlaGlnSer 1056
|||
Db 4229 TCAATGTCAGTGTCCAGGCT 4249
|||

RESULT 11
US-08-583-276-18
; Sequence 18, Application US/08583276
; Patent No. 5837536
; GENERAL INFORMATION:
; APPLICANT: McDonagh, Kevin T.
; APPLICANT: Nienhuis, Arthur
; APPLICANT: Tolstoshev, Paul
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HUMAN
; TITLE OF INVENTION: MULTIDRUG RESISTANCE GENES AND IMPROVED
; TITLE OF INVENTION: SELECTION OF CELLS TRANSFECTED WITH SUCH GENES
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan,
; ADDRESSEE: Cecchi & Stewart
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: DM4.V2
; CURRENT APPLICATION NUMBER: US/08/583,276
; FILING DATE: 05-JAN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/332,444
; FILING DATE: 31-OCT-1994
; APPLICATION NUMBER: 07/887,712
; FILING DATE: 22-MAY-1992
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4669 bases
; TYPE: nucleic acid
; STRANDEDNESS: singular
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: Genomic DNA
US-08-583-276-18
```

Alignment Scores:

```
Pred. No.: 0 Length: 4669
Score: 3056.50 Matches: 588
Percent Similarity: 74.86% Conservative: 210
Best Local Similarity: 55.18% Mismatches: 225
Query Match: 57.35% Indels: 43
DB: 2 Gaps: 6

US-09-873-409-4 (1-1058) x US-08-583-276-18 (1-4669)

QY 4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluVal 23
Db 1136 TCATTTACTGATTAAGAACTCTTAGCGGTATGCAAAAGCTGGACAGCTAGCTGAAGAGTGC 1195
|||
QY 24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43
|||
Db 1196 TTGGCAGCAATTAAGAACTGTGATTCATTTGGAGGACAAAGAAAGAACTTGA 1249
|||
QY 44 PheLeuLeuAsnIleThrArgTyrAlaTyrPheTyrPheProGlnTrpLeuLeuSerCys 63
Db 1249 ----- 1249
|||
QY 64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
Db 1250 -----AGGTACAAACAAAAATTTAGAAAGAGCTAAAAAGAAATTTGGATAAAG 1294
|||
QY 84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
Db 1295 AAAGCTATTACGCCAATATTTCTATAGGTGCTGCTTCTCGTGTATCTATGATCTAT 1354
|||
QY 104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyThrIle 123
Db 1355 GCTCTGGCCTCTGTTATGGACCACTTGGTCTCTCTCAGGGGAA-----TATTCTATT 1408
|||
QY 124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
Db 1409 GGACAAAGTACTCCTCTGTTATTTCTGTATTAATTTGGGGCTTTTAGTGTGGCAGGCA 1468
|||
QY 144 ValProHisPheGluThrPheAlaIleAlaArgGlyValAlaPheHisIlePheGlnVal 163
Db 1469 TCTCCAGCAATTGAAGCAATTTGCAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1528
|||
QY 164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
Db 1529 ATTGATAATAAGCCAAAGTATTGACAGCTATTGAGAGTGGGCACAAACACAGATAATATT 1588
|||
QY 184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
Db 1589 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATTCGAAAAGAAAGTTAAG 1648
|||
QY 204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
Db 1649 ATCTTGAGGGCCTGAACCTGAAGGTGCAGAGTGGCAGAGCGTCCCTCCCTGGTTGGAAC 1708
|||
QY 224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243
Db 1709 AGTGGCTGTGGGAAGAGCACACAGTCCAGCTGATGCAGAGGCTCTATGACCCACAGAG 1768
|||
QY 244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
Db 1769 GGGATGCTCAGTGTTCATGTCAGCAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 1828
|||
QY 264 HistLeuGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283
Db 1829 ATCATTTGGTGTGTAGTCAGGAACCTGTATTGTTGCCACCATGATGCTGAAACATTT 1888
|||
QY 284 LysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303
Db 1889 CGCTATGGCCGTGAATGTCACCATGGATGAGATGAGAAAGCTGTCAAGGAAGCCAAAT 1948
|||
QY 304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323
Db 1949 GCCTATGACTTTATCATGAAAGTGCCTCATAAATTTTGACACCTCTGTTGGAGAGAGAGGG 2008
|||
```


Qy	670	GlulLeuYlshisalaGlyIysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal	689
Db	3089	GAACATGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTCCGAACCGTTGTT	3148
Qy	690	SerLeuThrArgGluIysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis	709
Db	3149	TCCTTTGACTCAGAGACAGAAGTTTGAAACATATGTATGCTCAGAGTTTCAGGTACCATAC	3208
Qy	710	ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe	729
Db	3209	AGAAACTCTTTGAGGAAGACACACATCTTTTGGAAATTTACATTTTCTTACCACGCAATG	3268
Qy	730	IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg	749
Db	3269	ATGTATTTTTCCTATGCTGATGTTTCCGGTTTGAGAGCCTACTTGTGTGGCACAATAACTA	3328
Qy	750	MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle	769
Db	3329	ATGAGCTTTGAGATGTTCTGTGTAGTATTTTTCAGCTGTGTCTTTTGGTGCCATGCCCGT	3388
Qy	770	GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu	789
Db	3389	GGGCAAGTGAGTTCAITTTGCTCTGCTGATGATGCCAAGCCAAAATATCAGAGCCACATC	3448
Qy	790	PheAlaLeuLeuGluLysIysProAsnIleAspSerArgSerGlnGluGlyLysPro	809
Db	3449	ATCATGATCATTTGAAAAAACCCTTTTGATTGTACAGCTACAGCACGGAAGGCTAATGCCG	3508
Qy	810	AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro	829
Db	3509	AACACATTTGGAAGAAATGTCAATTTGTGTGAAGTTGATTCAACTATCTCCACCCCGACCG	3568
Qy	830	AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe	849
Db	3569	GACATCCAGTCCTTCAGGACTGAGCTTGGAGGTGAAGAGGGCCAGACGCTGGCTCTG	3628
Qy	850	ValGlySerSerGlyCysGlyIysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp	869
Db	3629	GTGGGCAGCAGTGGCTGTGGGAAGAGCAGATGGTCCAGCTCTCTGGAGCGGTCTTACGAC	3688
Qy	870	ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnThr	889
Db	3689	CCCTTTGCGAGGAAGTGTCTGTGTGTCGAAGAAATAAAGCGACTGAATTTTCAGTGG	3748
Qy	890	LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla	909
Db	3749	CTCGCAGCACACCTGGGCATCTGTCTCCAGAGAGCCCATCTCTGTTCATCGCAGCATTGCT	3808
Qy	910	GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla	929
Db	3809	GAGAACATTCCTATGACACACACAGCCGGTGGTGTACAGGAAGAGATCGTAGGGCA	3868
Qy	930	AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln	949
Db	3869	GCAAAGAGGCCAACATACATGCTTCATCTGAGTCACTGCTCTTAATAATATAGACACTAAA	3928
Qy	950	ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg	969
Db	3929	GTAGGAGACAAAGAACTCAGCTCTCTGTGTGCCAGAAACAACCGATTTGCCATAGCTCGT	3988
Qy	970	AlaLeuLeuGlnIysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn	989
Db	3989	GCCCTTGTAGACAGCCTCATATTTTGTCTTTTGGATGAAGCCACGTCAGCTCTGGATACA	4048
Qy	990	AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu	1009
Db	4049	GAAGAGTGAAGAGTTGTCCAAGAAGCCCTTGGACAAACCCAGAGAGGCGCGCACCTGCAAT	4108
Qy	1010	ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValLeuHisAsn	1029
Db	4109	GTGATGTCTCACCGCTGTCCACATCCAGAAATCAGACTTAATAGTGGTGTTCAGAAAT	4168
Qy	1030	GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe	1049

4169 GGATGAGTCAAGCAGCATGGCAGCATCAGCATGTCTGGCAGAGAGGCACTATTT 4228

Qy 1050 LysLeuValAsnAlaGln 1055
:::||||:::
pb 4229 TCAATGGTCAGTGTCCAG 4246

D_b 4229 TCAATGGTCAGTGTCAG 4246

RESIST 12

```

US-08-461-823-1
; Sequence 1, Application US/08461823
; Patent No. 5593840
; GENERAL INFORMATION:
; APPLICANT: Bhatnagar, Satish K.
; APPLICANT: George Jr., Albert L.
; APPLICANT: Nazarenko, Irina
; TITLE OF INVENTION: AMPLIFICATION OF NUCLEIC ACID SEQUENCES

```

Alignment Scores:					
Pred. No.:	3.5e-246	Length:	2726		
Score:	2280.00	Matches:	435		
Percent Similarity:	76.5%	Conservative:	159		
Best Local Similarity:	56.0%	Mismatches:	168		
Query Match:	42.78%	Indels:	14		
DB:	1	Gaps:	4		

ИТС-09-973-409-4 (1-1058) * ИТС-08-461-823-1 (1-2726)

Oy 295 MetGluArgAlaAalaArgGluAlaAsnAlaTyAspPheIleMetGluPheProAsnLys 314
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
nb 2 ATTGAGAAGTGTCAAGGAAGCAATGCTATGACTTTATCATGAACTCGCTCATAAA 61

ATTGAGAAAGCTGTCAAGGAAGCCAA TGCCTATGACTTTATCATGAAACTGCCTCATAAA 61

Qy 315 PheAsnThrLeuValGlyGluLysGlyAlaClnMetSerGlyGlyGlnLysGlnArgile 334

Db	62	TTTGACACCCCTGGTTGGAGAGAGAGGGCCCGAGTTGAGTGGTGGCGAGAACGACAGGATC	121
Qy	335	AlaIleAlaArgAlaLeuValArgAsnProIysIleLeuIleLeuAspGluAlaThrSer	354
Db	122	GCCATTGCACGTGCCCTGGTTGCAACCCCAAGATCCCTCTGCTGGATGAGGCCACGTCA	181
Qy	355	AlaLeuAspSerGluSerIysSerAlaValGlnAlaAlaLeuGluIysAlaSerIysGly	374
Db	182	GCCTTGGACACAGAAACGAGCAGCGTTCAGGTGGCTCTGGATTAAGGCCAGAAAGGT	241
Qy	375	ArgThrThrIleValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleVal	394
Db	242	CGACACCACTTGTGTAGCTCATCGTTGTCTACAGTTCGTAAATGCTGACGTCACTCG	301
Qy	395	ThrLeuIysAspGlyMetLeuAlaGluIysGlyAlaHisAlaGluLeuMetAlaIysArg	414
Db	302	GGTTTCGATGATGGAGTCATTGTGGAGAAAGAAATCATGATGAACCTCATGAAGAGAAA	361
Qy	415	GlyLeuTyrTyrSerLeuValMetSerGln-----AspIleIysIys	428
Db	362	GGCATTTTACTTCAACACTGTCAATGTCAGACAGCAGGAAATGAAGTTGAATTAGAAAT	421
Qy	429	AlaAspGluGlnMetGluSerMetThrTyrThrGluArgIysThrAsnSerLeuPro	448
Db	422	GCAGCTAGATGCCAAAGGTGAATGATGCTTGGAAATGCTTTCAAATGATTCAGA	481
Qy	449	LeuHisSerVal-----LysSerIleIys---SerAspPheIleAsp	461
Db	482	TCCAGTCTAATAGAAAAAGATCAACTCGTAGGAGTGTCGGTGGATCACAAGCCCAAGAC	541
Qy	462	LysAlaGluGluSerThrGlnSerIysGluIleSerLeuProGluValSerLeuLeuIys	481
Db	542	AGAAAGCTTAGTACCAAGAGGCGCTCGGATGAAGTATACCTCCAGTTTTCCTTTTGGAGG	601
Qy	482	IleLeuIysLeuAsnIysProGluTyrProPheValValLeuGlyThrLeuAlaSerVal	501
Db	602	ATTATGAAGCTAAATTTAACTGAATGGCCCTATTTTGTTGGTGTATTTTGTGCCATT	661
Qy	502	LeuAsnGlyThrValHisProValPheSerIleIlePheAlaIysIleIleThrMetPhe	521
Db	662	ATAAATGGAGCCCTGCACACAGCATTTGCCAATAATATTTTCAAAGATTATAGGGGTTTT	721
Qy	522	GlyAsn---AsnAspIysThrThrLeuIysHisAspAlaGluIleIysSerMetIlePhe	540
Db	722	ACAGAATTGATGATCTCGTAAACAAACACAGACAGATAGTAACCTGTGTTTCACTATGTGT	781
Qy	541	ValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArg	560
Db	782	CTAGCCCTTGGAAATTTATCTTTATTACATTTTTTCCCTCAGGGTTTCATTTTGGCAAA	841
Qy	561	AlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheIysAlaMetLeuTyrGln	580
Db	842	GCTGGAGAGATCCTCACCAAGCGCTCCGATACATGGTTTCCGATCCCATGCTCAGACAG	901
Qy	581	AspIleAlaTrpPheAspGluIysGluAsnSerThrGlyGlyLeuThrThrIleLeuIle	600
Db	902	GATGTGAGTGTGGTTGATGACCCCTAAAAACACCACTGGAGCATGACTACCAAGGCTCGCC	961
Qy	601	IleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsn	620
Db	962	AATGATGCTGCTCAAGTTTAAAGGGGCTTATAGGTTCCAGGCTTGCTGTAAATACCCAGAA	1021
Qy	621	AlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyrGlnMetThrPhe	640
Db	1022	ATAGCAAAATCTTGGACAGGAATTAATTAATCTTTCATCTATGGTTGGCACTAACACTG	1081
Qy	641	LeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMet	660
Db	1082	TTACTCTTAGCAATTGTAACCATCATCTTTCGAATAGCAGGAGTCTTTCGAAATGAAATGTTG	1141
Qy	661	ThrGlyPheAlaAsnIysAspGlyGlnGluLeuIysHisAlaGlyIysIleAlaThrGlu	680
Db	1142	CTTGGACAAGACACTGAAAGATTAAGAAAGACTGAAAGGTCCTGGGAAGATCGCTACTGAA	1201

Db 936 ----- 936
 QY 62 SerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 81
 :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
 Db 937 ----- -CAGTATGAGGTCCACTTTAGACGAAGCTTGAGAAATGGGGA 975
 QY 82 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGly 101
 |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
 Db 976 ACAAGAACAAGATGTCATGGGTTTCATGATTGGCGCATGTTTTGGCTTATGTACTCG 1035
 QY 102 ThrTyGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 121
 |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
 Db 1036 AACTACGGCTCTTGGCTTCGGATGGGTCTCTGTTCTTCGGTAGATGGT-----GCATGC 1089
 QY 122 ThriLeGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 141
 |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
 Db 1090 GATGTGGGTCAATAATCTCACAGTTCTCATGGCCACTTGTATCGGATCGTTCTCTCTGGGG 1149
 QY 142 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 161
 |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
 Db 1150 AACGTAGTCAAATGCTCAAGCATTTACAAACGCTGTGGCGCGGCCCAAGATATTT 1209
 QY 162 GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 181
 |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
 Db 1210 GGAAACGATCATCGCCAGTCCCATTAGATCATTAATCGAACGAAGGAAGACGCTCGAC 1269
 QY 182 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 201
 |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
 Db 1270 CATTTTGAGGGCCACATTTGATGCTACGCATGTCAAGCATATTTACCCTCTAGACCCGAG 1329
 QY 202 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 221
 :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
 Db 1330 GTCAACGCTCATGTGAGGATGTTCTCTGTCAATGTCGCGCTGGAAAAACAACGCTTTAGTC 1389
 QY 222 GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro 241
 |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
 Db 1390 GGGCCCTCTGGCTCTGGA AAAAGTAGTCGGTGGTCTGGTTGGAGCGAATCTCAGATCGCT 1449
 QY 242 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 261
 |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
 Db 1450 GTTCGGGTACGGTTTGTGGATGGCCATGACATCAGGACCTCAATCTCGTGGCTT 1509
 QY 262 ArgAspHisIleGlyValValSerGlnProValLeuPheGlyThrThrIleSerAsn 281
 |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
 Db 1510 CGCCAACAGATCTCTTTGGTTAGCCAGGAGCCTGTTCTTTTGGCACGACGATTATAAG 1569
 QY 282 AsnIleLysTyrGly-----ArgAspAspValThrAsp 292
 |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
 Db 1570 AATATTAGGCACGGTCTCATCGCAAAAAGTACGAGAATGAATCCGAGGATAAGCTCCGG 1629
 QY 293 GluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPhePro 312
 |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
 Db 1630 GAATCTATCAGAACCGCGCAAAATAGGGGAATGCTCATGACTTTATTACTGCCTTCGCT 1689
 QY 313 AsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGln 332
 |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
 Db 1690 GAAGTTATGAGCAATGTTGGCAGCGTGGCTTCTCCTTTTCCAGTGGCCAGAAACAG 1749
 QY 333 ArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAla 352
 |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
 Db 1750 CGCATTCGAATCGCCGTCGGTGTGTAGTGAACCCAAAAAATCTCTCTCTGGATGAAGCT 1809
 QY 353 ThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSer 372
 |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
 Db 1810 ACTTCGGCCTTGGACACAAATCCGAAGCGGTGGTTTCAAGCAGCTTTGGAGAGGCGAGCT 1869
 QY 373 LysGlyArgThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeu 392
 |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
 Db 1870 GAAGCCGAACACTACTATTGTGATCGCTCATCGCTTTTCCAGATCAAAACGGCCACAAC 1929
 QY 393 IleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAla 412
 |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
 Db 1930 ATTGTGTTCTGGTCAATGGCAAAATTTGCTGAACAGGAACCTCAGATCAATTTGTTGAC 1989

Qy	413	LyA	ArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGln	432
Db	1990	CGCGAGGCGCTTATCGAAACTTTGTGGAGGCTCAACGTATCAATGAACAGAGAAGACGT	2049	
Qy	433	MetGluSerMetThrTyrSerThrGluArgLysThrAsnSer	450	
Db	2050	GAGCCTTGGAGAGCGCCGACGCTGAGGATCTCAGCAATGAGATATTTGCCAAATCAAA	2109	
Qy	451	SerValLysSerIleLysSerAspPhe	462	
Db	2110	ACTGCGTCAAGCGCATCATCCGATCTCGACGGAAACCCACAAACCATTTGACCGCACGGC	2169	
Qy	463	AlaGluSerThrGlnSerLysGluIleSer	476	
Db	2170	ACCCACAAAGTCTGTTTCCAGCGCGATCTCTTCTAAAGACCCCGGAAACAACTCCGAAA	2229	
Qy	477	ValSerLeuLeuLysIleLeuLys	492	
Db	2230	TACTCATATGAGCGCTGCTCAAAATTTGTTGTTCTTCAACCGCCCTGAATCCCGTAC	2289	
Qy	493	ValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIle	512	
Db	2290	ATGCTCATCGGTCTTCTCTCTCAGTGTGTAGCTGTGTGGCCAAACCCACGCAAGACGATG	2349	
Qy	513	IlePheAlaLysIleIleThrMetPheGly	530	
Db	2350	CTATATGCTTAAGCAATAGCACACTCTCGTCCAGCAATCACATATATAGCAAGCTTCGA	2409	
Qy	531	HisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSer	550	
Db	2410	CATGATCGGATTTCTGGTCAATGATGTTCTTCGTGGTGTGTATCATTCAGTTTATCACG	2469	
Qy	551	TyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArg	570	
Db	2470	CAGTCAACCAATGGTGCTGCATTTGGCGTATGCTCCGAGAGACATTATTCGTGCGCGGAGA	2529	
Qy	571	HisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGluLysGluAsn	590	
Db	2530	AGCACTGCCCTTTCGGACGATACCTCCGTCAAGACATTTGCTTTTGACAAGGAAGAGAAT	2589	
Qy	591	SerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThr	610	
Db	2590	AGCACCGGCGCTCTGACCTCTTTCTCTGTCTCACCGAGAGCAACATCTCTCCGGTGTAGC	2649	
Qy	611	GlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIle	630	
Db	2650	GGTGTGACTCTAGGCAGATCTTGATGACCTCCAGACCCCTAGGAGCGGCTATCATTAAT	2709	
Qy	631	SerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeuAla	650	
Db	2710	GCCCTGGCGATTGGGTGGAAATGGCCCTTAGTTTGTATCTCGTTGTGCCGGTCTCTCCTG	2769	
Qy	651	ValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGlu	670	
Db	2770	GCATGCGGTTTCTACCGATCTTATATGCTACGCCAGTTTCAATACGCTCCAGCTTGCT	2829	
Qy	671	LeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSer	690	
Db	2830	TATGAGGATCTGCAAACTTTGCTTTCGAGGCTACATCGTCTATCCGCACAGTTGCGTCA	2889	
Qy	691	LeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArg	710	
Db	2890	TTAACCGGGAAAGGATGCTCTGGGAGATTTACCAATGCCAGCTTTCAGCCACAGGCGAGG	2949	
Qy	711	AsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIle	730	
Db	2950	ACCAGTCTAATCTCTGCTTGAGGTCAATCCCTGTATATGCGTCTCGACGACCTTGTT	3009	
Qy	731	TyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMet	750	
Db	3010	TTCTTCGCTGCTCGGTTTGGTACGGAGGACACTTTCTGGTGCACCAACAGATAT	3069	


```
QY 751 ThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGly 770
DB 3070 GACATTTCCTCGCTCTCTTTGTTGTTCTCGAGATTCTCTTGGTGTCTCAATCCGCGGC 3129
QY 771 LysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPhe 790
DB 3130 ACCGCTTTTCTTGCACACAGATGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3189
QY 791 AlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAsp 810
DB 3190 CGACTGTTCCAGCAAGGAGCACAATTTGATACTGCTGTAAGAGGAGGAGGAGGAGGAG 3249
QY 811 ThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAsp 830
DB 3250 ACCGTTGGAAGGTGAATTTAGGAACGTCACCTTCAGATACCCGAGCCGCGCCAGAA 3309
QY 831 ValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheVal 850
DB 3310 CAGCCTGTCTCGCGCGCTTGGACCTGACCGTGAACCTGGACAATATGTTGGCGTTGTC 3369
QY 851 GlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspPro 870
DB 3370 GGACCCAGCGGTGTGGCAAGATACCACTTGCATTGCTTGAGCGCTTTTACGATGCG 3429
QY 871 ValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGlnLeuAsnValGlnTrpLeu 890
DB 3430 ATTCCGCGGTCCATCTCTGTTGATGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3489
QY 891 ArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGlu 910
DB 3490 CGCAGCTTTCTGTCATCTGTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3549
QY 911 AsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAla 930
DB 3550 AACATCTTACTTGGTATTGTCGAAGATGACGTACCGAGAAATCTTGATTAAGCTTGC 3609
QY 931 AsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnVal 950
DB 3610 AAGGAGCGCTAATATCTACGACTTCTATCATGCTCGCTCCCGAGGCGCTTAAATACAGTTGT 3669
QY 951 GlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAla 970
DB 3670 GGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3729
QY 971 LeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAsp 990
DB 3730 CTCTTCGGGATCCCAAAATCTTCTTCGATGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3789
QY 991 SerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuVal 1010
DB 3790 TCAGAAAAGGTCTCCAGGCGGCTTTGGATGCGCTGCCCGAGGCGGAGCAACAATCGCC 3849
QY 1011 ValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGly 1030
DB 3850 GTTGCACACCGACTACGACGATTCAAAAGCGGAGCGTTATCTATGTTTTCGACCAAGGC 3909
QY 1031 LysIleLysGluGlnGlyThrHisGlnGluLeuArgAsnArgAspIleTyrPheLys 1050
DB 3910 AAGATCTCGAAAGCGGAGCGACGACGAGAACTGTCAGAAAAGGCGCGTACTACGAG 3969
QY 1051 LeuValAsnAlaGlnSerVal 1057
DB 3970 CTGGTCAACTTGCAGAGCTTG 3990
```

RESULT 14

```
US-08-996-545-3
; Sequence 3, Application US/08996545
; Patent No. 5928898
; GENERAL INFORMATION:
; APPLICANT: Skatrud, Paul L.
; APPLICANT: de Waard, Maarten A.
; APPLICANT: Peery, Robert B.
; APPLICANT: Andrade, Alan C.
```

```
; TITLE OF INVENTION: Multiple Drug Resistance Gene attrd of
; NUMBER OF INVENTIONS: Aspergillus nidulans
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: U.S.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/996,545
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Webster, Thomas D.
; REGISTRATION NUMBER: 39,872
; REFERENCE/DOCKET NUMBER: X-11766
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3334
; TELEFAX: 317-276-2763
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4002 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-996-545-3

Alignment Scores:
Pred. No.: 6,76e-222 Length: 4002
Score: 2067.00 Matches: 447
Percent Similarity: 59.43% Conservative: 199
Best Local Similarity: 41.12% Mismatches: 381
Query Match: 38.78% Indels: 60
DB: 2 Gaps: 9

US-09-873-409-4 (1-1058) x US-08-996-545-3 (1-4002)
QY 2 ValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGlu 21
DB 817 AUCAUCAAGUACAGCAAAAGUCGUAGACGACGAGGCGGCGGACGUGUUGCGGAA 876
QY 22 GluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnLysGluLeuGln 41
DB 877 GAGGUCACGACUCCAUCAAGAAUUGCCAGGCUUGGCGGCCACCCAGACAGCUUGCGAAG 936
QY 42 ArgSerPheLeuLeuAsnIleThrArgTyrAlaTyrPheTyrPheProGlnTrpLeuLeu 61
DB 936 ----- 936
QY 62 SerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysaspPheGly 81
DB 937 -----CAGUAUGAGGUCCACUAGACGAGCUGAGAAUUGGGA 975
QY 82 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGly 101
DB 976 ACNAGAACAGAUUGUUGGUGUUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUG 1035
QY 102 ThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 121
DB 1036 AACUACGUGUUGGUGUUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUG 1089
QY 122 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 141
DB ----- 141
```


Db 1090 GAUGUGGGGAUUAUUCACAGUUCUUCAGUGGCCCAUCUUGAUCGGAUGCUUUCUGGGG 1149
Qy 142 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 161
Db 1150 AACGUUAGCCAAAGCUCAGCAUUAUACAAACGCUGGCGCGCGCAAGAAUUAU 1209
Qy 162 GlnValIleAspLysPheProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 181
Db 1210 GGAACGAUCCAGCGCAGUCCCAUUAUAGUCCAUUAUUGAAACGAAGGAGAGCGUCGAC 1269
Qy 182 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 201
Db 1270 CAUUUGAGGGCCCAUUGAUAUACGCAUGCAAGCAUUAUUAUACCAUUAUAGCCGAG 1329
Qy 202 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 221
Db 1330 GUCACCGUGAGGAGGAUUGUUCUGCAUAGCCCGCGGAAACCAACCGCUUAGUC 1389
Qy 222 GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro 241
Db 1390 GGCCCUUCUGGUCUGGAAAGUACGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGG 1449
Qy 242 AspAspGlyPheIleMetValAspAsnAspIleArgAlaLeuAsnValArgHisTyr 261
Db 1450 GUUGCGGUAACGUUUUGGAGUGGCGCAUGACAAAGGACCUCAUCCUGCGUGCUU 1509
Qy 262 ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn 281
Db 1510 CGCCACAGAUUCUUGGUUAGCCAGAGCGUGUUCUUUUGGCGAGCAUUAUUAAG 1569
Qy 282 AsnIleLysTyrGly-----ArgAspAspValThrAsp 292
Db 1570 AAUAUAGGCAGCGUCUUCGCGCACAAAGUACGAGAAUAGCCGAGGAUAGGUGCCG 1629
Qy 293 GluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPhePro 312
Db 1630 GAACUCUACGAGAACCGCGGCAAAAGCGGAAUGCGCAUUAUUAUUAUUAUUAUUAU 1689
Qy 313 AsnLysPheAsnThrLeuValGlyLysGlyAlaGlnMetSerGlyGlyGlnLysGln 332
Db 1690 GAAGUUAUGAGACAAUUGUGGACGCGUGGUUUCUUCUUCAGUGGCGCAGAAACAG 1749
Qy 333 ArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAla 352
Db 1750 CGCAUUGCAUCCGCGCGUGGUGUUAUGACGCCAAUUAUUAUUAUUAUUAUUAUUAU 1809
Qy 353 ThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSer 372
Db 1810 ACUUGCGCCUUGGACACAAAUCCGAAUGCGGUGGUCAAGCAGCUUUGGAGAGGCGAGCU 1869
Qy 373 LysGlyArgThrThrIleValAlaHisArgLeuSerThrIleArgSerAlaAspLeu 392
Db 1870 GAAGCGCAACUUAUUGAUGCGUCAUGCCUUCUUCAGCAUCAAACCGCGCACAC 1929
Qy 393 IleValThrLeuLysAspGlyMetLeuAlaGlyLysGlyAlaHisAlaGluLeuMetAla 412
Db 1930 AUUGUGGUUUGGCAUUGGCAAAUUGUGAACAAGAACUCACGCAUUAUUAUUAUUAU 1989
Qy 413 LysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGln 432
Db 1990 CGCGGAGCGCUUAUUGCAAAUUGUGGAGGCGUCAAGCAUUAUUAUUAUUAUUAUUAU 2049
Qy 433 MetGluSerMetThrTyrSerThrGluArgLysThrAsnSer-----LeuProLeuHis 450
Db 2050 GACGCUUGGAGGAGCGCGAGCGUGAGAUUCUACGAUUGCAGAUUAUUGCCAAAUCAA 2109
Qy 451 SerValLysSerIleLysSerAspPhe-----IleAspLys----- 462
Db 2110 ACUGGCUAACGCGCAUCCAGUUCGACGCGAAACCCCAACCAUUAUUAUUAUUAUUAU 2169
Qy 463 AlaGluGluSerThrGlnSerLysGluIleSer-----LeuProGlu 476
Db 2170 ACCCACAGUCUGUUCAGCGCGAUUCUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 2229

Qy 477 ValSerLeuLeuLysIleLeuLys-----LeuAsnLysProGluTrpProPhe 492
Db 2230 UACUCAUAUUGAGCGCUGCUCAAAUUGUUGUUCUUCUACCGCCUGAAAUCGCGUAC 2289
Qy 493 ValValLeuGlyThrLeuAlaSerValLeuAenGlyThrValHisProValPheSerIle 512
Db 2290 AUGUCAUGCGGUGUGUUCUCAGUGUAGUGUGGUGGCGCAACCCACGCGAGAGUG 2349
Qy 513 IlePheAlaLysIleIleThrMetPheGly-----AsnAsnAspLysThrThrLeuLys 530
Db 2350 CUUAUGCUUAAGCCCAUCAGCACACUCUGCCGCCAGAAUACAUAUAGCAAGCUUCGA 2409
Qy 531 HisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSer 550
Db 2410 CAUGAUGCGGAUUCUGGCAUUGAUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGU 2469
Qy 551 TyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArg 570
Db 2470 CAGUCAACCAUUGGUGCUGCAUUGGCGUAUGCUCGAGAGACUAUUCGUCGCGGAGA 2529
Qy 571 HisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGluLysGluAsn 590
Db 2530 AGCACUGCCUUGCGGAGCAUACUCCGCAAGACAUGUUCUUAUUAUUAUUAUUAUUAU 2589
Qy 591 SerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThr 610
Db 2590 AGCACCGCGCGCUGACCUUUCUGUCCACGAGCAAGCAUCUCUCGUGUUGUAGC 2649
Qy 611 GlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIle 630
Db 2650 GGUGUGACUCUAGGCGACGAUUGAUGACUCCACACCCUAGGCGCGCAUUAUUAU 2709
Qy 631 SerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeuAla 650
Db 2710 GCCUGCGCAUUGGUGGAAUUGGCUUAUUGUUAUUGUUGUUGUUGUUGUUGUUGUUG 2769
Qy 651 ValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGlu 670
Db 2770 GCAUGCGGUUUCUACGUAUUAUUGUAGCCGCAUUAUUAUUAUUAUUAUUAUUAU 2829
Qy 671 LeuLysHisAlaGlyLysIleAlaThrGluAlaLeuLysAsnIleArgThrIleValSer 690
Db 2830 UAUGAGGGAUGCAACAAUUCUUGCGAGGCUAUCGUCUUAUUAUUAUUAUUAUUAUUA 2889
Qy 691 LeuThrArgGluLysAlaPheGluGlnMetTyrGluMetLeuGlnThrGlnHisArg 710
Db 2890 UUAACCGGGAAGGAGUGUGUGGAGAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 2949
Qy 711 AsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIle 730
Db 2950 ACCAGCUAAUUCUUGUUGAGGUCAUCCGUUAUUAUUAUUAUUAUUAUUAUUAUUAU 3009
Qy 731 TyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrIleIleGlnAlaGlyArgMet 750
Db 3010 UUCUUCUGGUGUGCGUGGUGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 3069
Qy 751 ThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGly 770
Db 3070 GACAUUUUCGCUUUCUUGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 3129
Qy 771 LysThrIleuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPhe 790
Db 3130 ACCGCUUUUUCUUGCACCAGAGGCGCAAGGCGGAAUUAUUAUUAUUAUUAUUAUUA 3189
Qy 791 AlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAsp 810
Db 3190 CGAGUGUUCACGGAACCAACAAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 3249
Qy 811 ThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAsp 830
Db 3250 ACGGUGGAGGAGGAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 3309

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 31, 2003, 13:52:37 ; Search time 385.104 Seconds
(without alignments)
2337.888 Million cell updates/sec

Title: US-09-873-409-4

Perfect score: 5330
Sequence: 1 WVISLTSKLSAYSKAGAVA.....QELLNRDIYFKLVNAQSVQ 1058

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 574371 seqs, 425486471 residues

Total number of hits satisfying chosen parameters: 1148742

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -DEV=xlh

-MODEL=frame+ p2n model -DEV=xlh
-Q=/cgn2_1/USPTO spool/US09873409/runat_27032003_115424_19290/app_query.fasta_1.7544
-DB=Published Applications NA -QMTX=fascap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09873409 @CGN 1.1 960 @runat_27032003_115424_19290
-NCPU=6 -ICPU=3 -NO XLPXY -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	5328	100.0	3177	10	US-09-873-409-12
2	5328	100.0	3702	10	US-09-873-409-13
3	5159.5	96.8	3621	10	US-09-873-409-14
4	4561.5	85.6	2856	10	US-09-873-409-10

5	3456	64.8	2066	10	US-09-873-409-9	Sequence 9, Appli
6	3111.5	58.4	4317	9	US-10-044-671-1	Sequence 1, Appli
7	3106.5	58.3	4369	10	US-09-769-097-1	Sequence 1, Appli
8	3106.5	58.3	4425	10	US-09-769-097-3	Sequence 3, Appli
9	3100.5	58.2	3860	10	US-09-866-866A-1	Sequence 1, Appli
10	3100.5	58.2	3860	10	US-09-866-866A-3	Sequence 3, Appli
11	3100.5	58.2	8630	10	US-09-306-417-1	Sequence 1, Appli
12	3100.5	58.2	8630	10	US-09-306-417-2	Sequence 2, Appli
13	3090.5	58.0	4788	10	US-09-866-866A-7	Sequence 7, Appli
14	3085	57.9	3924	10	US-09-880-107-2299	Sequence 2299, Ap
15	3085	57.9	4189	10	US-09-866-866A-5	Sequence 5, Appli
16	3084	57.9	4643	9	US-10-072-621-2	Sequence 2, Appli
17	3044.5	57.1	4254	10	US-09-917-800A-1424	Sequence 1424, Ap
18	3043	57.1	3912	10	US-09-917-800A-1560	Sequence 1560, Ap
19	2488.5	46.7	5010	10	US-09-917-800A-483	Sequence 483, App
20	2103	39.5	4175	10	US-09-749-340-3	Sequence 3, Appli
21	2072.5	38.9	3861	9	US-09-938-842A-2237	Sequence 2237, Ap
22	2067	38.8	4002	9	US-09-758-828-1	Sequence 1, Appli
23	2067	38.8	4002	9	US-09-758-828-3	Sequence 3, Appli
24	2039.5	38.3	3512	10	US-09-749-340-7	Sequence 7, Appli
25	1951	36.6	3861	9	US-09-938-842A-263	Sequence 263, App
26	1868	35.0	2021	10	US-09-873-409-15	Sequence 15, Appli
27	1806.5	33.9	4653	9	US-10-101-388-2	Sequence 2, Appli
28	1698.5	31.9	1940	10	US-09-873-409-16	Sequence 16, Appli
29	1617	30.3	2698	10	US-09-749-340-5	Sequence 5, Appli
30	1538.5	28.9	2681	10	US-09-749-340-8	Sequence 8, Appli
31	1460.5	27.4	3792	9	US-09-882-694-10	Sequence 10, Appli
32	1380	25.9	8036	9	US-10-101-388-1	Sequence 1, Appli
33	1326	24.9	3999	9	US-09-882-694-9	Sequence 9, Appli
34	1101.5	20.7	1175	10	US-09-873-409-11	Sequence 11, Appli
35	1032.5	19.4	7296	10	US-09-070-927A-59	Sequence 59, Appli
36	1002	18.8	9047	10	US-09-070-927A-12	Sequence 12, Appli
37	984.5	17.7	3741	9	US-09-738-626-1251	Sequence 1251, Ap
38	943.5	17.7	12438	10	US-09-070-927A-173	Sequence 173, App
39	902.5	16.9	3601	10	US-09-070-927A-255	Sequence 255, App
40	901.5	16.9	6415	10	US-09-070-927A-275	Sequence 275, App
41	868.5	16.3	640681	10	US-09-790-988-1	Sequence 1, Appli
42	836.5	15.7	1810	10	US-09-749-340-4	Sequence 4, Appli
43	758.5	14.2	2298	9	US-10-156-239-18	Sequence 18, Appli
44	758.5	14.2	2298	10	US-09-795-693-18	Sequence 18, Appli
45	758.5	14.2	3408	9	US-10-156-239-15	Sequence 16, Appli

ALIGNMENTS

RESULT 1
US-09-873-409-12
; Sequence 12, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
; APPLICANT: Frank, Markus
; APPLICANT: Sayegh, Mohamed
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873,409
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 3177
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: No. US20020037522A1e
; LOCATION: (198)..(198)
; OTHER INFORMATION: n at position 198 represents any nucleotide (A, T, C or G)
US-09-873-409-12

Alignment Scores:
Pred. No.: 0
Score: 5328.00
Percent Similarity: 100.00%
Length: 3177
Matches: 1058
Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.96% Indels: 0
DB: 10 Gaps: 0

US-09-873-409-4 (1-1058) x US-09-873-409-12 (1-3177)

```
Qy 1 MetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAla 20
Db 1 ATGGTCATCTCATTTGACCAAGTAAGATTAAGTGCCTATTCCAAAGCTGGGCTGTGCA 60
Qy 21 GluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeu 40
Db 61 GAAGAAGTCTTGTCATCAATCCGAACAGTCATAGCCTTTAGGGCCAGGAGAAAGAACTT 120
Qy 41 GlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeu 60
Db 121 CAAAGGCTCTTCCCTTTAAATATACAGATATGCTGGTGTATTTATTTCCCCAGTGGCTA 180
Qy 61 LeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPhe 80
Db 181 CTAAGTTGTGTTCTGTTNTTTGTAAGGTATACACAGATCTCAAGATGCAAGGATTTT 240
Qy 81 GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsn 100
Db 241 GGCATAAAAGGACTATAGCTTCAAAAGTGTCTCTGGTGTGTACTTCTTTATGAAT 300
Qy 101 GlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGly 120
Db 301 GGAACCTATGGACTTGCTTTTGTGTATGGAACCTCTTGATTTCTTAATGAGAACCTGGA 360
Qy 121 TyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIle 140
Db 361 TATACCATCGGACTGTCTTGCTGTTTCTTTAGTGTATCCATAGCAGTTATGCAAT 420
Qy 141 GlyAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 160
Db 421 GGAGCAGCAGTCCCTCACTTTGAAACCTTCGCAATAGCCCGAGGAGCTGCCTTTCAAT 480
Qy 161 PheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysPro 180
Db 481 TTCAGGTTATTTGATAGAAACCCAGATATAGATAACTTTTCCAGCTGGATATAAACCT 540
Qy 181 GluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgPro 200
Db 541 GAATCCATAGAGGACTGTGGATTTAAATATGTTCTTTCAATTCATCAAGACCA 600
Qy 201 SerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu 220
Db 601 TCTATCAAGATTCTGAAAGGTCTGAATCTCAGAATTAACTCTGGAGAGACAGTCGCCTG 660
Qy 221 ValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAsp 240
Db 661 GTCCGCTCTCAATGGCAGTGGGAAGAGTACGGTATGTCAGCTTCTGCAGAGGTTATATGAT 720
Qy 241 ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis 260
Db 721 CCGGATGATGGCTTTATCATGGTGGATGAGATGACATCAGAGCTTTAAATGTCGGCAT 780
Qy 261 TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSer 280
Db 781 TATCGAGACCATATTTGGAGTGGTTAGTCAAGAGCCTGTTTGTTCGGGACCCACCATCAGT 840
Qy 281 AsnAsnIleLysTyrGlyArgAspValThrAspGluLeuMetGluArgAlaAlaArg 300
Db 841 AACAAATCAAGATGACGAGATGATGTGATCTGATGAAGAGATGGAGAGACGCAAGG 900
Qy 301 GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly 320
Db 901 GAACAAATGCGTATGATTTTATCATGGAGTTTCTTAATAAATTTAATACATGCTAGG 960
Qy 321 GluLysGlyAlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu 340
Db 961 GAAAAAGAGAGCTCAAAATCAGTGGAGGGCAGAAACAGAGGATCGCAATTTGCTCGTGCCTTA 1020
```

```
Qy 341 ValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 360
Db 1021 GTTCGAAACCCCAAGATTTCTGATTTTAGATGAGGCTACGCTCTGCCCTGGATTTCAGAAAGC 1080
Qy 361 LysSerAlaValGlnAlaLeuGluLysAlaSerLysGlyArgThrThrIleValVal 380
Db 1081 AAGTCAGCTGTTCAAGCTGCTGAGAGAGCGGAGCAAGGTCGAGTACCAATCGTGGTA 1140
Qy 381 AlaHisArgLeuSerThrIleArgSerAlaAspIleValThrLeuLysAspGlyMet 400
Db 1141 GCACACCGACTTTCTACTATTCGAAGTGCAGATTTGATGTCACCTTAAAGGATGGAATG 1200
Qy 401 LeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeu 420
Db 1201 CTGGCGGAGAAAGAGCAGCATGCTGCACTAATGCAAAACAGAGTCTATATATTCACTT 1260
Qy 421 ValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThr 440
Db 1261 GTGATGTCACAGGATATTAATAAGCTGATGAACAGATGGAGTCAATGACATATTTCTACT 1320
Qy 441 GluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle 460
Db 1321 GAAAGAAAGACCAACTCACTTCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTTCATT 1380
Qy 461 AspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeu 480
Db 1381 GACAGCTGAGGAATCCACCAATCTAAAGAGATAGTCTTCTGAAAGTCTCTCTATTTA 1440
Qy 481 LysIleLeuLysLeuAsnLysProGluTrpPheValValLeuGlyThrLeuAlaSer 500
Db 1441 AAAATTTTAAAGTTAAACAAGCCTGAATGCCCTTTGTGCTTCTGGGAGCATTTGGCTTCT 1500
Qy 501 ValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMet 520
Db 1501 GTTCTAAATGGAACTGTTCATCCAGTATTTTCCATCATCTTTGCAAAAAATTAACCATG 1560
Qy 521 PheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePhe 540
Db 1561 TTTGGAAATTAATGATTAACCAACCATTAAGCATGATGCAGAAATTTATTCATGATATTC 1620
Qy 541 ValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArg 560
Db 1621 GTCATTTTGGCTGTATTTGCTTTTGTTCAGTTATTTTCATGCAGGATTTATTTACGGCAGA 1680
Qy 561 AlaGlyGluIleLeuThrMetArgLeuArgHisIleuAlaPheLysAlaMetLeuTyrGln 580
Db 1681 GCAGGGGAAATTTTAAAGATGAGATTAAAGACACTTGGCCTTCAAAAGCCATGTTATATCAG 1740
Qy 581 AspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAla 600
Db 1741 GATATTCCTGGTTTGTGTAAGAAAGGAAACAGCAGGAGGCTTGACAAATATTATGACC 1800
Qy 601 IleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsn 620
Db 1801 ATGATATATACAAATTTCAAGGAGCAACAGGTTCCAGGATTTGGCGTCTTTAAACACAAAAT 1860
Qy 621 AlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPhe 640
Db 1861 GCAACTAACATGGGACTTTTCAGTTATCTTCTTTATATATATGATGGAGATGACATTC 1920
Qy 641 LeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMet 660
Db 1921 CTGATTCAGTATTTGCTCCAGTACTTGGCGTGACAGAAATGATTTGAAACCGCAGCAATG 1980
Qy 661 ThrGlyPheAlaAsnLysAspLysGlnLeuLysHisAlaGlyLysIleAlaThrGlu 680
Db 1981 ACTGGATTTCCCAACAAAGATAAGCAAGAACTTAAAGCATCTCGAAAGATAGCAACTGAA 2040
Qy 681 AlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMet 700
Db 2041 GCTTTGGAGAAATATACGTACTATAGTGTCTATTAAACAGGGGAAAAAGCCTTCGAGCAAAATG 2100
```



```
QY 701 TyrGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGly 720
Db 2101 TATGAAGAGATGCTTACAGATCTACACAGAAATACCTCGAAGAAGCACAGATTTATTGGA 2160
QY 721 SerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPhe 740
Db 2161 AGCTGTATGCATTCACCAATGCCCTTTATATATTTTCCCTATGCAGCAGGGTTTCGATTT 2220
QY 741 GlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThr 760
Db 2221 GGAGCCTATTATTCAGCTGGACGAATGACCCACAGAGGCGATGTTTCATAGTTTTCAT 2280
QY 761 AlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSer 780
Db 2281 GCAATTGCATATGGAGCTATGGCCATCGGAAAAACGCTCGTTTGGCTCTCGAATATTC 2340
QY 781 LysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAsp 800
Db 2341 AAGCCAAATCGGGGCTGGCATCTGTTTGCCTTGTGGNAAAGAAACCAATATAGAC 2400
QY 801 SerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGlu 820
Db 2401 AGCGCAGTCAAGAAGGCAAAAGCCAGACACATGTCAAGGGAATTTAGAGTTTCGAGAA 2460
QY 821 ValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSer 840
Db 2461 GTCTCTTCTTCTATCCATGTCGCCAGATGTTTTCATCTCCGTGGCTTATCCCTCAGT 2520
QY 841 IleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSer 860
Db 2521 ATTGAGCGAAGAAAGACAGTAGCATTGTGGGAGACGGCTGTGGGAAAGCACTTCT 2580
QY 861 ValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyVal 880
Db 2581 GTTCAACTTCTGCAGAGACTTATGACCCCGTGCAGACAAAGTCTGTTGATGGNGTG 2640
QY 881 AspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGlu 900
Db 2641 GATGCAAAAGAAATTAATGTATACAGTGGCTCGGTTCCCAAAATAGCAATCGTTCTCAAGAG 2700
QY 901 ProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnArgVal 920
Db 2701 CCGTGTCTTCAACTCGAGCATTTGTGAGAACATCGCTATGATGACACAGCCGGTGT 2760
QY 921 ValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGlu 940
Db 2761 GTGCCATTAGATGAGATCAAGAAGCCGCAATGACCAATATCCATTTCTTTATTGAA 2820
QY 941 GlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGly 960
Db 2821 GGTCTCCCTGAGAAATACACACACAAGTTGGACTGAAAGGAGCACAGCTTTCTGGCGGC 2880
QY 961 GlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeu 980
Db 2881 CAGAAACAAAGACTAGTATTGTCAAGGGCTCTTCTCCAAAACCCAAAATTTTATTGTTG 2940
QY 981 AspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAsp 1000
Db 2941 GATGAGCCCACTTCAGCCCTCGAATATGACAGTGAAGAGTGGTTTCAGCATGCCCTTGAT 3000
QY 1001 LysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsn 1020
Db 3001 AAAGCCAGGACGGGAAGGACATGCTAGTGTGCTCACTCACAGGCTCTCTGCAATTCAGAAC 3060
QY 1021 AlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlu 1040
Db 3061 GCAGATTGATAGTGGTTCTGCAAAATGGAAGATTAAGAAACCAAGAACTCATCAAGAG 3120
QY 1041 LeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 1058
Db 3121 CTCTGGAATAATCGAGACATATATTTTAAGTTAGTAGTGAATGACAGTCAAGTGCAG 3174
```

RESULT 2

```
US-09-873-409-13
; Sequence 13, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
; APPLICANT: Frank, Markus
; APPLICANT: Sayegh, Mohamed
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873,409
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 13
; LENGTH: 3702
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: No. US20020037522A1e
; LOCATION: (723)..(723)
; OTHER INFORMATION: n at position 723 represents any nucleotide (A, T, C or G)
US-09-873-409-13
```

```
Alignment Scores:
Pred. No.: 0 Length: 3702
Score: 5328.00 Matches: 1058
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99,96% Indels: 0
DB: 10 Gaps: 0
```

US-09-873-409-4 (1-1058) x US-09-873-409-13 (1-3702)

```
QY 1 MetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAla 20
Db 526 ATGGTCATCTCATGTGACAGTAGGAATTAAGTGCCTATTCGAAGCTGGGGCTGTGGCA 585
QY 21 GluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeu 40
Db 586 GAAGAAGTCTTGTCATCAATCCGACAGTCATAGCTTTAGGGCCCGAGGAGAAAGAACTT 645
QY 41 GlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeu 60
Db 646 CAAAGGTCTTTCCTTTTAAATATAACAAGATATGCTGGTGTATTTTCCCAAGTGGCTA 705
QY 61 LeuSerCysValLeu**PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPhe 80
Db 706 CTAAAGTGTGTTCTGTTNTTTGTAAGGTATACAGAACTCAAGATGCAAGGATTTT 765
QY 81 GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsn 100
Db 766 GGCATAAAAGGACTATAGCTTCAAAAGTGTCTCTTGGTGTCTGTACTTCTTTATGAAT 825
QY 101 GlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerIleLeuAsnGlyGluProGly 120
Db 826 GGAACCTATGGAATCTGCTTTTGGTATGGAACCTCTGATTTCTTAATGGAACCTGGA 885
QY 121 TyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIle 140
Db 886 TATACCATCGGAGCTGTCTTGTCTGTTTCTTTAGTGTATTCATGACGATTTATGCATT 945
QY 141 GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 160
Db 946 GGAGCAGCAGTCCCTCACITTTGAAACCTTCGCAATAGCCCGAGGAGCTGCTTTCATATT 1005
QY 161 PheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysPro 180
Db 1006 TTCAGGTATTGATAAGAAACCCAGTATAGATAAATCTTTTCCACAGCTGATATAAACCT 1065
QY 181 GluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgPro 200
Db 1066 GAATTCATAGAAGGAATCTGGAAATTTAAAAATGTTTCTTTCATATTATTCATCAAGACCA 1125
```


QY 201 SerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu 220
Db 1126 TCTATCAAGATTCTGAAGGCTCAATCTCAGAAATTAACTCTGGAGAGACAGTCGCCTTG 1185
QY 221 ValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAsp 240
Db 1186 GTCCGCTCTCAATGGCAGTGGGAAGAGTACGGTAGCTCCAGCTTCTGCGAGAGGTATATGAT 1245
QY 241 ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis 260
Db 1246 CCGGATGATGCTTTATCATGGTGGATGAGAAATGACATCAGAGCTTTAAATGTGGCGCAT 1305
QY 261 TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSer 280
Db 1306 TATCGAGCCATATTGGAGTGGTTAGTCAAGAGCCTGTTTGTTCGGGACCACCATCAGT 1365
QY 281 AsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArg 300
Db 1366 AACAAATATCAAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAGAGCAGCAAGG 1425
QY 301 GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly 320
Db 1426 GAAGCAAAATGCGTATGATTTTATCATGGAGTTTCTTAATAAATTTAATACATTCGTGAGG 1485
QY 321 GluLysGlyAlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu 340
Db 1486 GAAAAAGGAGCTCAATCAGTGGAGGCGCAGAAACAGAGGATCGCAATTCGTGCGCTTA 1545
QY 341 ValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 360
Db 1546 GTTCGAAACCCCAAGATTCTGATTTTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGC 1605
QY 361 LysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValVal 380
Db 1606 AAGTCAGCTGTTCAGCTGCACTGGAGGCGCAGAAAGTCGGACTACAAATCGTGGA 1665
QY 381 AlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMet 400
Db 1666 GCACCCGACTTCTACTATTTCGAATGCAGATTTGATTTGTGACCCTAAGGATGGAATG 1725
QY 401 LeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeu 420
Db 1726 CTGCGGAGAAAGGAGCACATGCTGAATATGCAATATGCAAAACAGAGTCTATATTTACT 1785
QY 421 ValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThr 440
Db 1786 GTGATGTACAGGATATTAATAAAGCTGATGAACAGATGGAGTCAATGACATATTTACT 1845
QY 441 GluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle 460
Db 1846 GAAAGAAGACCAACTCACTTCTCTGCACTCTGTGAAGACATCAAGTCAGACTTCATT 1905
QY 461 AspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeu 480
Db 1906 GACAAGGCTGAGGAATCACCCCAATCTAAAGAGATAAGTCTTCTGGAAGTCTCTCTATTA 1965
QY 481 LysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSer 500
Db 1966 AAAATTTTAAAGTTAAACAAGCCCTGAATGGCCTTTTGTGGTTCCTGGGACATTCGCCTCT 2025
QY 501 ValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMet 520
Db 2026 GTTCTAAATGAACTGTTTCATCCAGTATTTTCCATCATCTTTGCAAAAATTAATACCATG 2085
QY 521 PheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePhe 540
Db 2086 TTTGCAAAATATGATAAACACATTAAGCATGATGACAGAAATTTATTCATCATATTC 2145
QY 541 ValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArg 560
Db 2146 GTCATTTGGGTGTATTTGCTGATGATTTTTCATGACAGGATATTTTACGGCAGA 2205
QY 561 AlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGln 580

Db 2206 GCAGGGGAAATTTTAAACGATGAGATTAAAGACACTTGGCCCTTCAAGCCCATGTTATATCAG 2265
QY 581 AspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAla 600
Db 2266 GATATTCCCTTGGTTTGTATGAAAAGGAAACACAGCACAGAGGCTTGAACAATATTAGCC 2325
QY 601 IleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsn 620
Db 2326 ATAGATATAGCAAAATTCAGGAGCAACAGGTTCCAGGATTCGCGCTTTAAACACAAAAT 2385
QY 621 AlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPhe 640
Db 2386 GCAACTAACATGGGACTTTTCAGTTATCATTTCTTTATATATGATGGAGAGATGACATTC 2445
QY 641 LeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMet 660
Db 2446 CTGATTCGTAGTATTGCTCCAGTACTTGGCGTGACAGGAATGATTAAGAACCGCAGCAATG 2505
QY 661 ThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGlu 680
Db 2506 ACTGGATTTGCCAACCAAGNATAGCAAGAACTTAAGCATGCTGGAAGATAGCACTGAA 2565
QY 681 AlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMet 700
Db 2566 GCTTTGGAGAAATATACGTACTATAGTGTCAATTAACAAGGGAAAAAGCCTTCGAGCAAAATG 2625
QY 701 TyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGly 720
Db 2626 TATGAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCACAGATATTGGA 2685
QY 721 SerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPhe 740
Db 2686 AGCTGTATGATTCAGCCATGCTTTATATATATTTTGCCTATCGACAGGGTTTCGATTT 2745
QY 741 GlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThr 760
Db 2746 GGAGCCTATTTAATCAAGCTGAGCAATGACCCAGAGGGCATGTTTCATAGTTTACT 2805
QY 761 AlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSer 780
Db 2806 GCAATTCATATGGAGCTATGGCCATCGGAAAAACGCTCGTTTGGCTCTCGAATATTC 2865
QY 781 LysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuGluLysLysProAsnIleAsp 800
Db 2866 AAGCCAAATCGGGGCTGCGCATCTGTTTGCCTTGTGGAAAAAGAAACCAATATAGAC 2925
QY 801 SerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGlu 820
Db 2926 AGCCCGCAGTCAAGAAGGAAAAAACCCAGACACATGTAAGGGAATTTAGAGTTTCGAGAA 2985
QY 821 ValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSer 840
Db 2986 GTCTCTTCTTCTCATCATGTCGCCAGATGTTTTCATCTCCGTGGCTTATCCCTCAGT 3045
QY 841 IleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSer 860
Db 3046 ATTGAGCGAGAAAGACAGTAGCATTTTGGGAGGACGCGGCTGTGGGAAAAAGCATTTCT 3105
QY 861 ValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyVal 880
Db 3106 GTTCAACTCTGCGAGAGACTTTATGACCCCGTCAAGGACAAAGTGTGTTGATGGTGTG 3165
QY 881 AspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGlu 900
Db 3166 GATGCAAAAAGAAATGAATGTACAGTGGCTCCGTTCCCAATAGCAATCGTTCCTCAAGAG 3225
QY 901 ProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgVal 920
Db 3226 CCTGTGCTTCTCACTGCGACATTTGTGAGAAACATCGCCTATGTTGACACACGCCGTGTG 3285
QY 921 ValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGlu 940


```
Db 3286 GTCCATTAGATGAGATCAAGACCGCAAAATGCAGCAATATCCATCTCTTTATTGAA 3345
Qy 941 GlyLeuProGluLysTyrAenThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGly 960
Db 3346 GGTCTCCCTCAGAAATACAAACACAACTGGGACTGAAAGGAGCACAGCTTTCTGCGCGC 3405
Qy 961 GlnLysGlnArgLeuAlaLeuAArgAlaLeuGlnLysProLysIleLeuLeu 980
Db 3406 CAGAAACAAAGACTAGCTATGCAAGGGCTCTCTCCAAAACCCCAAAATTTATTGTTG 3465
Qy 981 AspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAsp 1000
Db 3466 GATGAGGCCACTTCAGCCCTCGATATGACAGTGAAGAGTGGTTCCAGCATGCCCTTGAT 3525
Qy 1001 LysAlaArgThrGlyArgThrCysLysLeuValValThrHisArgLeuSerAlaIleGlnAsn 1020
Db 3526 AAAGCCAGGACGGGAAGGACATCGCTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAAC 3585
Qy 1021 AlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlu 1040
Db 3586 GCAGATTTGATAGTGGTTCTGCACAAATGGAAGATTAAGGAACAGGAACCTCATCAAGAG 3645
Qy 1041 LeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 1058
Db 3646 CTCTGAGAAATCGACATATATTTTAAGTTAGTGAATGCACAGTCAGTGCAG 3699

RESULT 3
US-09-873-409-14
; Sequence 14, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
; APPLICANT: Frank, Markus
; APPLICANT: Sayegh, Mohamed
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
; FILE OF INVENTION: Homologue on Chromosome 7p15-21 and Uses Thereof
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873,409
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
; LENGTH: 3621
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-873-409-14

Alignment Scores:
Pred. No.: 0 Length: 3621
Score: 5159.50 Matches: 1031
Percent Similarity: 97.45% Conservative: 0
Best Local Similarity: 97.45% Mismatches: 0
Query Match: 96.80% Indels: 27
DB: 10 Gaps: 1

US-09-873-409-4 (1-1058) x US-09-873-409-14 (1-3621)
Qy 1 MetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAla 20
Db 526 ATGGTCATCTCATGACAGTAAGAAATTAAGTCCCTATTCCTCAAGCTGGGGCTGTGGCA 595
Qy 21 GluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeu 40
Db 586 GAAGAAAGTCTTGTCATCAATCCGACAGTCATAGCCCTTTAGGGCCCGAGGAGAAAGAACTT 645
Qy 41 GlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTyrPheTyrPheProGlnTrpLeu 60
Db 646 CAA----- 648
Qy 61 LeuSerCysValLeu**PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPhe 80
Db 649 -----AGTATACACAGAAATCTCAAGATGCAAAAGGATTTT 694
Qy 81 GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsn 100
```

```
Db 685 GGATAAAAAGGACTATAGCTTCAAAAGTGTCTCTGGGTGCTGTGTACTTCTTTATGAAT 744
Qy 101 GlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGly 120
Db 745 GGAACCTATGAGACTTGTCTTTTGGTATGGAACCTCTCTGATCTCTTAATGGAGAACCTGGA 804
Qy 121 TyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIle 140
Db 805 TATACCATCGGAGCTGTCTCTGCTGTTTCTTTAGTGTAAATCCATAGCAGTTATTGCATT 864
Qy 141 GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 160
Db 865 GGAGCAGCAGTCCCTCCTTGTAAACCTTCGCAATAGCCGAGGAGCTGCTTTCATATT 924
Qy 161 PheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysPro 180
Db 925 TTCAGGTTATTGATAGAAACCCAGTATAGATAACTTTTCCACAGCTGGATATAAACCT 984
Qy 181 GluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgPro 200
Db 985 GAATCCATAGAAGAACTGTGGAATTTAAAAATGTTCTTTCAATATTCATCAAGACCA 1044
Qy 201 SerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu 220
Db 1045 TCTATCAAGATTTCTGAAGGTCTGAAATTTAAGTCTGGAGAGACAGTCGCTTGTG 1104
Qy 221 ValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAsp 240
Db 1105 GTCCGTCTCAATGGCAGTGGAGAGTACGTTAGTCCAGCTTCTGCAGAGTTTATATGAT 1164
Qy 241 ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis 260
Db 1165 CCGGATGATGGCTTTTATCATGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGCAT 1224
Qy 261 TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSer 280
Db 1225 TATCGAGACCATATTGGAGTGGTGTAGTCAAGAGCCCTGTTTGTTCGGGACCCATCAGT 1284
Qy 281 AsnAsnIleLysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArg 300
Db 1285 AACATATCAAGTATGGACAGATGATGTGACTGATGAAGATGGAGAGACGAGCAAGG 1344
Qy 301 GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly 320
Db 1345 GAAGCAATATGCGTATGATTTTATCATGGAGTTTCTCTAAATAATTTAATACATTTGGTAGG 1404
Qy 321 GluLysGlyAlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu 340
Db 1405 GAAAAGGAGAGCTCAATGAGTGGAGGCGAGAAACAGAGGATCGCAATTCGTCGCTTTA 1464
Qy 341 ValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 360
Db 1465 GTTCGAAACCCCAAGATTTCTGATTTTAGATGAGGCTTACCTCTGCCCTGGATTCAAGAAAGC 1524
Qy 361 LysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValVal 380
Db 1525 AAGTCAGCTGTTCAAGCTGCACCTGGAGAGGCGAGCAAAAGGTCGGAATCAATTCGTGTA 1584
Qy 381 AlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMet 400
Db 1585 GCACACCGACTTCTACTATTTCGAAGTGCAGATTTGATTTGTGACCCCTAAAGGATGGAATG 1644
Qy 401 LeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeu 420
Db 1645 CTGGCGGAGAAAGGAGCACATCTGAACTTAATGGCAAAACGAGGCTCTATATTATTCACCT 1704
Qy 421 ValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThr 440
Db 1705 GTGATGTCACAGATATTTAAAAAAGCTGATGAACAGATGAGTCAATGACATATTTCTACT 1764
Qy 441 GluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle 460
```


Db 1765 GAAAGAAAGACCAACTCCTCTCTGCACTCTGTGTGAAGAGCATCAAGTCAGACTTCATT 1824
QY 461 AspLysAlaGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeu 480
Db 1825 GACAAGGCTGAGGAATCCACCCTAATCTAAGAGATAGTCTTCTGAAGTCCTCTATT 1884
QY 481 LysIleLeuLysLeuAsnLysProGluTyrProPheValValLeuGlyThrLeuAlaSer 500
Db 1885 AAAATTTAAAGTTAAACAAGCTGAATGCCCTTTGTGGTCTCTGGGGACATTTGGCTTCT 1944
QY 501 ValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMet 520
Db 1945 GTTCTAAATGGAACTGTGTATCCAGTATTTTCCATCATCTTTGCAAAAATTAACCATG 2004
QY 521 PheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePhe 540
Db 2005 TTTGGAAATTAATGATATAAACACCATTAAGCATGATGACAGAAATTTATTCATGATATTC 2064
QY 541 ValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArg 560
Db 2065 GTCATTTTGGGTGTTATTTGCTTTGTCTGTTATTTTCATGCAGGGATTTATTTACGGCAGA 2124
QY 561 AlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGln 580
Db 2125 GCAGGGGAAATTTTAACGATGAGATTAAAGACACTTGGCCTTCAAGGCCATGTTATATCAG 2184
QY 581 AspIleAlaTyrPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAla 600
Db 2185 GATATTCCTCGTTTGTATGAAGAAAGAAACAGACAGAGGCTTGCACAAATATTAGCC 2244
QY 601 IleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsn 620
Db 2245 ATAGATATACACAAATTTCAAGGAGCAACAGGTTCCAGGATTTGGCGTTTAAACACAAAT 2304
QY 621 AlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyrGluMetThrPhe 640
Db 2305 GCAACTAACATGGGACTTTCAGTTATCATTTCTTTATATATGATGGAGATGACATTC 2364
QY 641 LeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMet 660
Db 2365 CTGATTTCTGAGTATTTGCTCCAGTACTTGCCTGACAGAAATGATTGAAACCGCAGCAATG 2424
QY 661 ThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGlu 680
Db 2425 ACTGGATTTGCCACAAAGATAACCAAGAACTTAAGCATGCTGGAAAGATAGCAACTGAA 2484
QY 681 AlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMet 700
Db 2485 GCTTTGGAGAATATACGTACTATATAGTGTCTATTAAACAAGGGGAAAAGGCTTCGAGCAAAATG 2544
QY 701 TyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGly 720
Db 2545 TATGAAGAGATGCTTTCAGACTCAACAGAAATACCTCGAAGAAAGCACAGATTTATTGGA 2604
QY 721 SerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPhe 740
Db 2605 AGCTGTTATGATTCAGCCATGCCCTTTATATATTTTCCCTATGCATGCAGGGTTTCGATTT 2664
QY 741 GlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThr 760
Db 2665 GGAGCCTATTAAATCAAGCTGGACGAATGACCCAGAGGCGATGTTTCATAGTTTCTTACT 2724
QY 761 AlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSer 780
Db 2725 GCAATTCATATGAGCTATGGCCATCGAAAAACCGTCGTTTGGCTCTCGAATATTTCC 2784
QY 781 LysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAsp 800
Db 2785 AAAGCCAAATCGGGGGTGGCCATCTGTTGCCCTTGTGGAAAAAGAAACCAATATAGAC 2844
QY 801 SerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGlu 820
Db 2845 AGCCGCGAGTCAAGAGGAAAAAGCCAGACATGTGAAGGGAATTTAGAGTTTCGAGAA 2904

QY 821 ValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSer 840
Db 2905 GTCTCTTTCTTATCCATGTGCGCCAGATGTTTTTCATCTCCGTGGCTTATCCCTCAGT 2964
QY 841 IleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSer 860
Db 2965 ATTGAGCGAGGAAGACAGTAGCATTTTGGGGAGCAGCGGCTGTGGGAAAGACATTTCT 3024
QY 861 ValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlnValGlnValLeuPheAspGlyVal 880
Db 3025 GTTCAACTTCTGCAGAGACTTTATGACCCCGTGCAGAGGACAGTGTCTTTGATGTGTG 3084
QY 881 AspAlaLysGluLeuAsnValGlnTyrLeuArgSerGlnIleAlaIleValProGlnGlu 900
Db 3085 GATGCAAAAAGAAATTGAATGTACAGTGGCTCCGTTCCCAATAGCAATCGTTCTCTCAAGAG 3144
QY 901 ProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgVal 920
Db 3145 CCTGTGCTCTTCAACTGCGAGCATTTGCTGAGAACATCGCCTATGTTGTCACACGCCGTGTG 3204
QY 921 ValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGlu 940
Db 3205 GTGCCAATTAGATGAGATCAAGAGCCGCAATGCGACAAATATCCATCTCTTTATTGAA 3264
QY 941 GlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGly 960
Db 3265 GGTCTCCCTGAGAAATACAAACACACAAGTTGGACTGAAAGGAGCACAGCTTTCTGGCGGC 3324
QY 961 GlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeu 980
Db 3325 CAGAAACAAAGACTAGCTATTGCAAGGGCTCTTCTCCAAAAACCCAAAAATTTTATTGTTG 3384
QY 981 AspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAsp 1000
Db 3385 GATGAGGCCACTTCAGCCCTCGAATAATGACATGAGAGGTGGTTTCAGCATGCGCTTGAT 3444
QY 1001 LysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsn 1020
Db 3445 AAAGCCAGGACGGAGGACATGCTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAAC 3504
QY 1021 AlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlu 1040
Db 3505 GCAGATTTGATAGTGTCTTCTGCACAATGGAAGATAAAGAACCAAGGAACTCATCAAGAG 3564
QY 1041 LeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 1058
Db 3565 CTCTGAGAAATCGACACATATATTTTAAGTTAGTGAATGCACAGTCAGTGCAG 3618

RESULT 4

US-09-873-409-10
; Sequence 10, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
; APPLICANT: Frank, Markus
; APPLICANT: Sayegh, Mohamed
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
; TITLE OF INVENTION: Homologue on Chromosome 7p15-21 and Uses Thereof
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873,409
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
; LENGTH: 2856
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-873-409-10

Alignment Scores:
Pred. No.: 0 Length: 2856
Score: 4561.50 Matches: 912
Percent Similarity: 97.96% Conservative: 0

Best Local Similarity: 97.96%		Matches: 1	
Query Match: 85.58%		Indels: 18	
DB:		Gaps: 1	
US-09-873-409-4 (1-1058) x US-09-873-409-10 (1-2856)			
QY	129	ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHis-PheGI	148
DB	114	GTATTCTTTAGTGAATCCATAGCAGTTATTGCATTTGGAGCAGCAGTCCCTCATTATT--	171
QY	148	uThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPr	168
DB	172	-----GATAGAAACC	182
QY	168	oSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGI	188
DB	183	CAGTATAGATAAATCTTTCCACAGCTGGATATAAACCTGAATCCATAGAAAGAACTGTGGA	242
QY	188	uPheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLe	208
DB	243	ATTATAAAATGTTTCTTTCAATTATCCATCAAGACCATCTATCAAGATTCTGAAAGGTCT	302
QY	208	uAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLysAsnGlySerGlyLY	228
DB	303	GAATCTCAGAAATTAAGTCTGGAGAGCAGTCGCTTGGTCTCAATGGCAGTGGGAA	362
QY	228	sSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVa	248
DB	363	GAGTACGGTAGTCCAGCTTCTGCAGAGTTATATGATCCGATGATGGCTTTATCATGGT	422
QY	248	lAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVa	268
DB	423	GGATGAGAAATGACATCAGAGCTTTAAATGTGCGGCATTATTCGAGACCATATTGGAGTGGT	482
QY	268	lSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAs	288
DB	483	TAGTCAAGAGCCCTGTTTGTTCGGGACCCATCAGTAAACAAATATCAAGTATGGACGAGA	542
QY	288	pAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheII	308
DB	543	TGATGTGACTGATGAGAGATGGAGAGCAGCAGCAGGAGCAAAATGCGTATGATTTAT	602
QY	308	eMetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGlyAlaGlnMetSerGI	328
DB	603	CATGGAGTTTCTTAATAAATTAATACATTTGTTAGGGGAAAAGGAGCTCAAAATGAGTGG	662
QY	328	yGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuII	348
DB	663	AGGGCAGAAAACAGAGGATCGCAATTCGTCTGCGCTTAGTTTCGAAACCCCAAGATTCGAT	722
QY	348	eLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLe	368
DB	723	TTTAGATGAGGCTACGCTCGCCCTGGATTTCAGAAAGCAAGTCAGCTGTTCAAGCTGCAC	782
QY	368	uGlyLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIleAr	388
DB	783	GGAGAAGCGGAGCAAAAGGTCCGACTACAATCGTGTGTAGCACACCGACTTTCTACTATTCG	842
QY	388	gSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAl	408
DB	843	AAAGTGCAGATTGATTGTGACCCCTAAAGGATGGAATGCTGCGGAGAAAGGAGCAGCATGC	902
QY	408	aGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIleLysLY	428
DB	903	TGAATCTAATGGCAAAACGAGGTCTATATTATTACCTTGTGTATGTCACAGGATATTAAAAA	962
QY	428	sAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuPr	448
DB	963	AGCTGATGAACAGATGGAGTCAATGACATATTCTACTGAAAGAAAGAACCACTCACTTCC	1022
QY	448	oLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThrGI	468
DB	1023	TCTGCACTCTGTGAAGACATCAAGTCAGACTTCATTGACAAGGCTGAGGAATCCACCCA	1082
QY	468	nSerLysGluIleSerLeuProGluValSerLeuLysIleLeuLysLeuAsnLysPr	488
DB	1083	ATCTAAAGAGATAAGTCTTCTCCTGAAGTCTCTCTATTAAAAAATTTAAAGTTAAACAAGCC	1142
QY	488	oGluTyrProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPr	508
DB	1143	TGAATGGCCCTTTGTGGTCTTGGGACATTTGGCTTCTGTTCTAAATGGAACTGTTTATCTC	1202
QY	508	oValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrTh	528
DB	1203	AGTATTTCATCATCTTTGCAAAAATTTAAACCATGTTTGGAAATATATGATAAACCCAC	1262
QY	528	rLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPh	548
DB	1263	ATTAAAGCATGATGCAGAAATTTATTCCATGATATTTCGTCAATTTGGTGTATTATTCCTT	1322
QY	548	eValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetAr	568
DB	1323	TGTCAGTTATTTCATGCAGGGATTATTTTACGGCAGAGCAGGGGAAATTTTAAACGATGAG	1382
QY	568	gLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGluLY	588
DB	1383	ATTAAAGCATCTTGGCTTCAAGCCATGTTATATCAGGATATTGCTTGGTTCATGNAAA	1442
QY	588	sGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGI	608
DB	1443	GGAAAAACAGCACAGGAGGCTTGACAACAATATTAGCCATAGATATAGCAAAATTTCAAGG	1502
QY	608	yAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerVa	628
DB	1503	AGCAACAGGTTCCAGGATTGGCGTCTTAAACAAAAATGCAACTTAACATGGGACTTTTCAGT	1562
QY	628	lIleIleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProVa	648
DB	1563	TATCATTTCTTTTATATGATGGATGGGAGATGACATTCCTGATTCCTGAGTATTCCTCAGT	1622
QY	648	lLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLY	668
DB	1623	ACTTGGCGTGACAGGAATGATTGAAACCCGACGCAATGACTGGATTTGCCAACAAAGATAA	1682
QY	668	sGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrII	688
DB	1683	GCAAGAAGCTTAAGCATGCTCGAAAGATAGCAACTGAAGCTTTTGGAGAATATACGTACTAT	1742
QY	688	eValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGI	708
DB	1743	AGTGTCAITTAACAAAGGAAAAGCCCTTCGAGCAAAATGTTATGAAGAGATGCTTCAGACTCA	1802
QY	708	nHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAI	728
DB	1803	ACACAGAAATACCTCGAAGAAAGCACAGATTAATTGGAAGCTGTTTATGCAITTCAGCCATGC	1862
QY	728	aPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGI	748
DB	1863	CTTTATATATTTTGGCTATGTCAGCAGGGTTTCGATTTGGAGCCCTATTTAATTTCAAGCTGG	1922
QY	748	yArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAl	768
DB	1923	ACGAATGACCCAGAGGGCATGTTTCATAGTTTTCGCAATTCGCAATTCGATATGGAGCTATGGC	1982
QY	768	aIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHi	788
DB	1983	CATCGAANAACCGCTCGTTTGGCTCTGAAATATTCCAAAGCCCAATTCGGGGCTCGGCA	2042
QY	788	sLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLY	808
DB	2043	TCTGTTTGGCTTGTGGAAAAAGAAACCAATATAGACAGCCGAGTCAAGAGGGGAAAAA	2102
QY	808	sProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysAr	828
DB	2103	GCCAGACACATGTGAAGGGAAATTTAGAGTTTCGAGAAGTCTCTTTCTTCTATCATCTGTCG	2162


```
QY 828 gProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAl 848
DB 2163 CCCAGATGTTTTCATCTCCGTCGCTTATCCCTCAGTATTGAGCGAGAAAGACAGTAGC 2222
QY 848 aPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTy 868
DB 2223 ATTTGTGGGAGCAGCGCGCTGTGGGAAAGCAGCTTCTGTTCACCTCTCTGAGAGACTTTA 2282
QY 868 rAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGlnLeuLeuValG1 888
DB 2283 TGACCCCGTGCAGGACAGTCTGTTGATGTGTGATGTGATGCAAAAGATTTGAATGTACA 2342
QY 888 nTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIl 908
DB 2343 GTGGCTCCGTTCCCAATAGCAATCGTTCTCTCAAGAGCGCTGTGCTCTTCAACTGCAGCAT 2402
QY 908 eAlaGluAsnIleAlaTyArgAspAsnSerArgValValProLeuAspGluIleLysG1 928
DB 2403 TGCTTGAGAACATCGCTATGTGACAAACAGCCGCTGTGTGTCATTAGATGAGATCAAGA 2462
QY 928 uAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyAsnTh 948
DB 2463 AGCCGCAATGCAGCAATATCCCAATCTTTTATTGAAGGCTCTCCCTGAGAAATACAACAC 2522
QY 948 rGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAl 968
DB 2523 ACAAGTTGGACTGAAGAGACACAGCTTCTGGCGGCGCAGAAACAAAGACTAGCTATTGC 2582
QY 968 aArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAs 988
DB 2583 AAGGCTCTTCTCCAAAACCCAAAATTTTATTGTGATGAGGCCACTTCAGCGCCCTCGA 2642
QY 988 pAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCy 1008
DB 2643 TAAATGCAGTGAAGAGGTGTTCAGCATGCCCCCTTGATAAAGCCAGGAGGAGAGCATG 2702
QY 1008 sLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHi 1028
DB 2703 CCTAGTGTCTACTCACAGCTCTCTGCAATTCAGAACGCGAGATTTGATGTTCTGCA 2762
QY 1028 sAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTy 1048
DB 2763 CAATGGAAGATAAAGGAACAAGGAACCTCATCAAGAGCTCCTGAGAAATCGAGACATATA 2822
QY 1048 rPheLysLeuValAsnAlaGlnSerValGln 1058
DB 2823 TTTTAAGTTAGTGAATGCACAGTCAGTGCAG 2853
```

RESULT 5

```
US-09-873-409-9
; Sequence 9, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
; APPLICANT: Frank, Markus
; APPLICANT: Savegh, Mohamed
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
; TITLE OF INVENTION: Homologue on Chromosome 7p15-21 and Uses Thereof
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873,409
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 2066
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-873-409-9
```

Alignment Scores:	
Pred. No.:	3,38e-314
Score:	3456.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Length:	2066
Matches:	687
Conservative:	0
Mismatches:	0

```
Query Match: 64.84% Indels: 0
DB: 10 Gaps: 0
US-09-873-409-4 (1-1058) x US-09-873-409-9 (1-2066)
QY 372 SerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAsp 391
DB 3 AGCAAGGTCGGAGTACAATCGTGGTAGCACACCGACTTTTCTACTATTCCGAAAGTCAGAT 62
QY 392 LeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMet 411
DB 63 TTGATTGTGACCTTAAGGATGGAATGCTGGCGAGAAAGGAGCACATGCTGAACATAATG 122
QY 412 AlaLysArgGlyLeuTyTrpSerLeuValMetSerGlnAspIleLysLysAlaAspGlu 431
DB 123 GCAAAACGAGGCTATATTATTACCTTGATGTACAGGATATTAAAAAGCTGATGAA 182
QY 432 GlnMetGluSerMetThrTyTrpSerThrGluArgLysThrAsnSerLeuProLeuHisSer 451
DB 183 CAGATGGAGTCAATGACATATTCTACTGAAAGAAAGACCAACTCCTTCTCTGCACTCT 242
QY 452 ValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThrGlnSerLysGlu 471
DB 243 GTGAAGAGCATCAAGTCAGACTTCATTGACAGGCTGAGGAATCCACCCCAATCTAAGAG 302
QY 472 IleSerLeuProGluValSerLeuLysIleLeuLysLeuAsnLysProGluTrpPro 491
DB 303 ATAAGTCTTCTCAAGTCTCTCTATTAAAAATTTTAAAGTTAAACCAAGCCTGAATGSCCT 362
QY 492 PheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSer 511
DB 363 TTTGTGGTCTCTGGGACATTTGGCTTCTGTTCTAAATGGAACTGTTCAATCCAGTATTTTCC 422
QY 512 IleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThrLeuLysHis 531
DB 423 ATCATCTTTGCCAAAATTTATTAACATGTTTGGAAATTAATGATAAACCACATTAAGCAT 482
QY 532 AspAlaGluIleTyTrpSerMetIlePheValIleLeuGlyValIleCysPheValSerTy 551
DB 483 GATGCAAAATTTATTTCATGATATTCGTCAATTTTGGGTGTTATTGCTTGTGTCAGTTAT 542
QY 552 PheMetGlnGlyLeuPheTyArgAlaGlyGluIleLeuThrMetArgLeuArgHis 571
DB 543 TTCATGAGGGATTTATTTTACGCGAGAGCGAGGAAATTTTAAAGATGAGATTAAGACAC 602
QY 572 LeuAlaPheLysAlaMetLeuTyTrpGlnAspIleAlaTrpPheAspGluLysGluAsnSer 591
DB 603 TTGGCTTCAAGCCATGTTATATCAGGATATTCCTGCTGGTTTGGTGAAGAAAGAAACAGC 662
QY 592 ThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGly 611
DB 663 ACAGGAGGCTTGACAAATATTAGCCATAGATATAGCAAAATTCAGGAGCAACAGGT 722
QY 612 SerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSer 631
DB 723 TCCAGGATTTGGCGTCTTAAACACAAAATGCAACTAACATGGGACTTTTTCAGTTATCATTTCC 782
QY 632 PheIleTyTrpGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaVal 651
DB 783 TTTATATATGATGGAGATGACATTCCTGATTCCTGAGTATGCTCCAGTACTTGGCGTG 842
QY 652 ThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeu 671
DB 843 ACAGGAATGATTGAACCGCAGCAATGACTGGAATTTGCCAACAAAGATAAGCAAGAACTT 902
QY 672 LysHisAlaGlyLysIleAlaThrGluAlaLeuLysAsnIleArgThrIleValSerLeu 691
DB 903 AAGCATGCTGGAAGATAGTCAACTGAAAGCTTTGAGAAATATACGTACTATATAGTGTCA 962
QY 692 ThrArgGluLysAlaPheGluGlnMetTyTrpGluGluMetLeuGlnThrGlnHisArgAsn 711
DB 963 ACAAGGGGAAAAGCCCTTCGAGCAAAATGTATGAGAGAGATGCTTTCAGACTCAACACAGAA 1022
```


Db 3391 CAGTGGCTCGAGCACACACCTGGGCATCGTCTCTCAGGAGCCCATCTGTTGACTGCAGC 3450
Qy 908 IleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLys 927
Db 3451 ATTGCCGAGAACATTGGCTATGAGAGCAACAGCCGGTCTATCAGTGAAGAGATTATG 3510
Qy 928 GluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluIlyTyrAsn 947
Db 3511 CAGGCGCCAGAGGAGGCAACATACACCTTTCATCGAGACACTCCCTCGAGAAATCAAC 3570
Qy 948 ThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIle 967
Db 3571 ACCAGAGTAGGACACAAAGAACCCAGCTCTCTGGTGGCCAGAAACAGCCATGGCCATA 3630
Qy 968 AlaArgAlaLeuLeuGlnTyrProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeu 987
Db 3631 GCTCGCGCTCTGTAGACAGCTCATATTTGCTTTTGGATGAAGCTACATCAGCTCTG 3690
Qy 988 AspAsnAspSerGluLysValGlnHisAlaLeuAspLysAlaArgThrGlyArgThr 1007
Db 3691 GATACAGAAAGTGAAAGGTGTCCAAAGAGCCCTGGACAAAGCCAGAGAAGCCGCCACC 3750
Qy 1008 CysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeu 1027
Db 3751 TGCAITGTGATGCCCCACCCTGTCACCATCCAGAAATGCAGATTTAATAGTGGTGT 3810
Qy 1028 HisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlnLeuLeuArgAspIle 1047
Db 3811 CAGAATGGCAAGTCAAGGAGCATGGCACACATCACAGCTGTGGCTCAGAAAGGCATC 3870
Qy 1048 TyrPheLysLeuValAsnAlaGlnSer 1056
Db 3871 TATTTTTCATGATCAGTGTCCAGGCT 3897

RESULT 7

US-09-769-097-1
; Sequence 1, Application US/09769097
; Patent No. US20020055128A1
; GENERAL INFORMATION:
; APPLICANT: Kimberly Anne Brun
; APPLICANT: Richard James Chenery
; APPLICANT: Harma Eilens
; APPLICANT: John Anthony Feild
; APPLICANT: Lin Yue
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES
; FILE REFERENCE: GP-50009-C2
; CURRENT APPLICATION NUMBER: US/09/769,097
; CURRENT FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 09/208,809
; PRIOR FILING DATE: 1998-12-09
; PRIOR APPLICATION NUMBER: 09/156,800
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: US99/20770
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 4369
; TYPE: DNA
; ORGANISM: RATTUS RATTUS
US-09-769-097-1

Alignment Scores:
Pred. No.: 5,7e-281 Length: 4369
Score: 3106.50 Matches: 600
Percent Similarity: 75.54% Conservative: 206
Best Local Similarity: 56.23% Mismatches: 218
Query Match: 58.28% Indels: 43
DB: 10 Gaps: 7

US-09-873-409-4 (1-1058) x US-09-769-097-1 (1-4369)

Qy 4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluVal 23
Db 1039 TCATTTTTCATGATAAGAACTCCAGGCTTATGCAAAAGCTGGAGCACTTCTGAAGAAGTC 1098
Qy 24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGlnLysGluLeuGlnArgSer 43
Db 1099 TTAGACGCATCAGAACTGTGATTGCTTTGGAGGACAAAGAGGAACTTGA----- 1152
Qy 44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63
Db 1152 ----- 1152
Qy 64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
Db 1153 -----AGGTACAAATAACAATTTGGAAGAGCTAAAGAGGTGGGATAAAG 1197
Qy 84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
Db 1198 AAGAGTATCAGGCCCAACATTTCCATGGGTGAGCTTTTCTGCTTATCTATGCATCATAT 1257
Qy 104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123
Db 1258 GCTCTGGCATTCGTGTATGGACTTCCTTGGTCATCTCAAAAGAA-----TACACTATT 1311
Qy 124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
Db 1312 GCACAAAGTCTCAGCTCTCTTTTCTGTATTAATTTGGAGCATTCAGTGTGGCAGGCA 1371
Qy 144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163
Db 1372 TCTCAAATATTGAAGCCTTCGCCAATGCTAGAGGAGCAGCTTATGAAGTCTTCAGTATA 1431
Qy 164 IleAspLysGlyProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
Db 1432 ATTGATAATAAGCCAGTATAGACAGCTTCTCAAGAGAGTGGGCACAAACCCGACACATA 1491
Qy 184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
Db 1492 CAAGGAAATTTGGAATTCAAAATATTTCATCTTCAAGTACCCGTCCTCGAAAAGACGTTCAG 1551
Qy 204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
Db 1552 ATCTTGAGGGCTCAACCTGAAGGTGAAGCGGCACACGCTAGCCCTGGTTGGCAAC 1611
Qy 224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243
Db 1612 AGTGGCTGTGGGAAAAGACAACTGTCCAGCTGTCTCAGAGGCTCTACGACCCCATAGAG 1671
Qy 244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
Db 1672 GCGAGGTCAATATCGACGACAGGACATCAGGACCATCAATGTGAGGTATCTGCGGAA 1731
Qy 264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283
Db 1732 ATCATTTGGGGTGTGAGTCAGGAACCCGCTGTGTTGCCACCAATTCGCCAAAACATT 1791
Qy 284 LysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303
Db 1792 CGCTATGGCCGAGAAAACGTCACCATGGATGATAGAGAAAGCTGTCAAGGAGGCAAT 1851
Qy 304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323
Db 1852 GCCTATGATTTTCATCATGAACTGCCCCCAAAATTTGACACCTCTGTTGTTGAGAGGG 1911
Qy 324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343
Db 1912 GCGCAGCTGAGTGGGGGACAGAAACAGAGATCGCCATTTGCCCGGGCCCTGGTCCGCAAC 1971
Qy 344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
Db 1972 CCCAAGATCTCTTTGTTGATGAGGCCAGCTCAGCTTGGACACAGAAAGCGGCGGTG 2031
Qy 364 ValGlnAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383

Db 2032 GTTCAGCGCGCTCTGGATAAGCTAGAGAGCGCGACCACTTGATGATGATCACCAGC 2091
Qy 384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403
Db 2092 TTGTCTACAGTTCGCAATGCTGACGTCATTTGCTGTTTGGTGTGTCATTTGGAG 2151
Qy 404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423
Db 2152 CAAGGAATCATGATGATGCTCATGAGAGAGAAATTTACTTCAAACTTGTCATGACT 2211
Qy 424 Gln-----AspIleLysAlaAspGluGlnMetGluSer----- 435
Db 2212 CAGACAGCAGGAATGAATGAATAGGAATCAAGCTTGTAATCTAAAGACGGAATT 2271
Qy 436 -----MetThrTyrSerThrGluArgLysThrAsnSerLeu-----ProLeuHisSer 451
Db 2272 GATAATGTGGACATGCTCTCAAAAGATTCGGGATCCAGCTTAATAAGAGAAGATCAACT 2331
Qy 452 ValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThr---GlnSerLys 470
Db 2332 CGCAAAAGCATCCGTGGGCCACATGATCAAGACGGGAACCTTAGCACCAAGAGGCTCTG 2391
Qy 471 GluIleSerLeuProGluValSerLeuLysIleLeuLysLeuAsnLysProGluTyr 490
Db 2392 GATCAGCAGCTACCTCCAGCTTCCTTTGGCGGATCCTGAAGTTGAATCAACTGAATGG 2451
Qy 491 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 510
Db 2452 CCTTATTTGTGGTGTGTTATTTGTGCCATAAATGAGGCTTGCAACCAAGCATTC 2511
Qy 511 SerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrLeu--- 529
Db 2512 TCCATAATATTTCAAGAGTTGTAGGGGTTTTTACAAAAATGACACCCCTGGAATCCAG 2571
Qy 530 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 549
Db 2572 CGGCAGAACACCACTGTTTCTTATTGTTCTGTATCTCTGGATCATCTCTTTCATT 2631
Qy 550 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 569
Db 2632 ACGTTTTCTTCAGGGCTTCACATTTTGGCAAAGCTGGAGAGATCCTCACCAAGCGACT 2691
Qy 570 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGluLysGlu 589
Db 2692 CGATCATGCTCTCAAAATCCTGTGAGACAGGACATTAAGCTGTTGATGACCTTAA 2751
Qy 590 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 609
Db 2752 AACACCACAGGAGCGCTGACCCACAGGCTTGCCAATGACGCTGCTCAAGTCAAAAGGGCT 2811
Qy 610 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629
Db 2812 ACAGGCTCTAGGCTTGCTGTTATTACCCAGAACATAGCAAAATCTTGGGACAGGATCATC 2871
Qy 630 IleSerPheIleTyrGlyTyrGluMetThrPheIleLeuLeuSerIleAlaProValLeu 649
Db 2872 ATATCCCTGATCTAGGCTGGCAATTTGACACTTTTACTCTAGCAATTTGCCATCAIT 2931
Qy 650 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 669
Db 2932 GCTATAGCAGGAGTGGTTGAATGAAATGTTGTCTGGACAGCGCTGAAAGATAAGAAG 2991
Qy 670 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 689
Db 2992 GAACTTAGAAGTTCTGGGAAGATCGCTACAGAAAGCAATTTGAAACTTTTCGCACTGTCG 3051
Qy 690 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 709
Db 3052 TCTTTGACTCGGAGCAGAAAGTTGAAGACTATGATGATGCCAGAGCTTCAGATACCATC 3111
Qy 710 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 729

Db 3112 AGAAATGCTTTGAGAAAGCGCAGCTCTTTGGGATCAGCTTTCTCTCCACCCAGGCCATG 3171
Qy 730 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyValaTyrLeuIleGlnAlaGlyArg 749
Db 3172 ATGATTTCTCCATCTGCTGCTTTCCGTTTGTGCTTACTTGTGGCACAGAACTC 3231
Qy 750 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 769
Db 3232 ATGACATTTGAAATGTTCTGTTAGTATTCTCAGCTATTGTTCTTTGGTGCATGGCAGTG 3291
Qy 770 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaHisLeu 789
Db 3292 GGGCAGCTGAGTTTCATTCGCTCTGACTAGCGAAAGCCAAAGCTCTCGGCATCCCCACAT 3351
Qy 790 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 809
Db 3352 ATCAGATCATTTGAGAAATCCCTGAGATTGACAGTACAGACAGGAGGCTTGAAGCCT 3411
Qy 810 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 829
Db 3412 AATATGTTGGAAGAAATGTGAAATTTAATGGAGTCATGTTCAACTATCCACCCGACCC 3471
Qy 830 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 849
Db 3472 AACATCCAGTGTCTCAGGGCTGAGCTAGAGGTGAAGAAAGGAGAGCGCTGGGCCCTC 3531
Qy 850 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 869
Db 3532 GTGGGAGCAGTGGCTGCGGGAAGAGTACAGTGTCCAGTGTCTTGAGCGCTTCTATGAC 3591
Qy 870 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTyr 889
Db 3592 CCCATGCCGAGACAGTGTCTTAGATGGCAAGAAATAAGCAACTCAATGCTCCAGTGG 3651
Qy 890 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909
Db 3652 CTCGGGCCACCTCGGCAATTTGTGCCAGAGCCCATCTCTGTTGATCGACATCGCC 3711
Qy 910 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 929
Db 3712 GAGAACATTCCTACGGAGACAAAGCCGTGCTGTCTCATAGGAGATCGTGAAGGCA 3771
Qy 930 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 949
Db 3772 GCCAAGGAGGCCAACATCCACCACTCATCGACTGCTGAGAAATACAAACCCAGA 3831
Qy 950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 969
Db 3832 GTGGAGACAAAGGGACTCAGCTGTGCGGCGGCGAGAGCGCATCGGCATCGCGCGC 3891
Qy 970 AlaLeuLeuGlnLysProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989
Db 3892 GCCCTGCTCAGACAGCTCATCTTACTTCTGATGAGCGACATCAGCTCTGGATACG 3951
Qy 990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009
Db 3952 GAGAGTCAAAAGGTCTGTCAGGAAGCGCTGCAAAAGCCAGGAGGCCCGCACCTGCATT 4011
Qy 1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1029
Db 4012 GTGATCGCAGCGCTGTCCACCATCCAGACGACAGTGTGATCGGTGATTCAGAAC 4071
Qy 1030 GlyLysIleLysGluGlnGlyThrHisGlnLeuLeuArgAsnArgAspIleTyrPhe 1049
Db 4072 GCCAGGTCAAGGAGACCGGACCCACAGAGCTGTGATCGGTGATTCAGAAC 4131
Qy 1050 LysLeuValAsnAlaGlnSer 1056
Db 4132 TCGATGTGTCAGTGGCAGGCT 4152
RESULT 8
US-09-769-097-3
; Sequence 3, Application US/09769097

Db 2572 CGCAGAACAGCAACTGTTCTTTCTTTATTTGTTCTGATCTCTGGATCATCTCTTTTCATT 2631
Qy 550 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluLeuLeuThrMetArgLeu 569
Db 2632 ACGTTTTCCTTCAGGGCTTCACATTTGGCAAGCTGGAGAGATCTTCACCAAGCGACTC 2691
Qy 570 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 589
Db 2692 CGATACATGGTCTTCAATCCATCTGAGACAGGACATAAGCTGGTTGGTATGACCTTAA 2751
Qy 590 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 609
Db 2752 AACACCAAGGAGGCTGACACAGGCTTGGCAATGACGCTGCTCAAGTGAAGGGGCT 2811
Qy 610 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629
Db 2812 ACAGGGTCTAGGCTTGTGTTATTACCCAGAACATAGCAATCTTGGGACAGGATCATC 2871
Qy 630 IleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeu 649
Db 2872 ATATCCCTGATCTACGGCTGGCAATTGACACTTTTACTCTAGCAATTGTTCCCATCAT 2931
Qy 650 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 669
Db 2932 GCTATAGCAGAGTGGTGAATGAAATGAAATGTTGTCTGACAAAGCGCTGAAGATAAG 2991
Qy 670 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 689
Db 2992 GAACTAGAAGTCTCGGAAGATCGCTACAGAACAATTGAAAACTTTCCGACACTGTCTG 3051
Qy 690 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 709
Db 3052 TCTTTGACTCGGAGCAGAAAGTTTGAAATGATGATGCCAGAGCTTGCAGATACCATC 3111
Qy 710 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 729
Db 3112 AGAAATGCTTTGAAGAAAGCGACAGCTGTTGGGATCATCTTCTCTTCACCCAGGCGCATG 3171
Qy 730 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 749
Db 3172 ATGATTTCTCCTATGCTGTTGTTTCCGGTTTATGCTGCTACTTGGTGGCAGGAATC 3231
Qy 750 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 769
Db 3232 ATGACATTTGAAATGTTCTCTAGTATCTCAGCTATTGCTTGGTGGCATGCGCATG 3291
Qy 770 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 789
Db 3292 GGGCAGGTTCAGTTTCATTCGCTCCTGACTACGCGAAAGCAAGTCTCGGCATCCACATC 3351
Qy 790 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 809
Db 3352 ATCAGGATCATTTGAAATAATCCCTGAGATTGACAGCTACAGCAGGAGGGCTTGAAGCCT 3411
Qy 810 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 829
Db 3412 AATATGTTGGAGGAATGTAATTAATGAGTTCATGTTCACTATCCACCCGACCC 3471
Qy 830 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 849
Db 3472 AACATCCAGTGTCTCAGGGCTGAGCCTAGAGGTGAAGAAAGGCGACAGCTGGCCCTC 3531
Qy 850 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuGlnArgLeuTyrAsp 869
Db 3532 GTGGCAGCAGTGGCTGCGGAAGATGACGTGGTCCAGCTGCTGAGCGCTTCTATGAC 3591
Qy 870 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 889
Db 3592 CCCATGGCGGAGACAGTGTCTTCTAGATGGCAAGAAATAAGCAACTCAATGTCAGTGG 3651
Qy 890 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909
Db 3652 CTCGCGCCCACTGGGCATTTGTGTCCAGGAGGCCATCTCTGTTTACTGTCAGCATGCC 3711

Qy 910 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluLeuLysGluAla 929
Db 3712 GAGAACATTCCTACGAGACAAACAGCCGTCGTCTCTATAAGGAGATCGTGAAGGCA 3771
Qy 930 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 949
Db 3772 GCCAAGGAGGCCAACATCCACAGTTTCATCGACTCATCTCCCTGAGAAATACCAACAC 3831
Qy 950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 969
Db 3832 GTGGAGACAAAGGACTCAGCTGTCGGCGGCGCAGACAGCGCATCGCCATCGCGCGC 3891
Qy 970 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAsn 989
Db 3892 GCCTCTGTCAGACAGCTCCTCATCTTTACTTCTTGATGAAGCAGCATCAGCTCTGGATACG 3951
Qy 990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009
Db 3952 GAGAGTGAAGAGTCTCCAGGAAGCGCTCGACAAAGCCAGGAGGCGCACCTGCATT 4011
Qy 1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1029
Db 4012 GTGATCGGACCGCTCTGCCACCATCCAGACGACACTTGATCGTGTGATTCAGAAC 4071
Qy 1030 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1049
Db 4072 GGCCAGGTCAAGGAGCAGCGCACCCACAGCAGCTCTGCCCCAGAAAGGACATCTATTTC 4131
Qy 1050 LysLeuValAsnAlaGlnSer 1056
Db 4132 TCGATGTCAGTGTGCAGGCT 4152

RESULT 9

US-09-866-866A-1
; Sequence 1, Application US/09866866A
; Patent No. US20020102244A1
; GENERAL INFORMATION:
; APPLICANT: Sorrentino, Brian
; APPLICANT: Schuetz, John
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
; FILE REFERENCE: 1340-1-021CIP2
; CURRENT APPLICATION NUMBER: US/09/866,866A
; CURRENT FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 09/584,586
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: PCT/US99/11825
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: 60/086,988
; PRIOR FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 3860
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-866-866A-1

Alignment Scores:
Pred. No.: 1,73e-280 Length: 3860
Score: 3100.50 Matches: 595
Percent Similarity: 75.54% Conservative: 211
Best Local Similarity: 55.76% Mismatches: 218
Query Match: 58.17% Indels: 43
DB: 10 Gaps: 6

US-09-873-409-4 (1-1058) x US-09-866-866A-1 (1-3860)

Qy 4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluVal 23
Db 712 TCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCTGAGCAGTAGCTGAAGAGTC 771
Qy 24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43

Db 1285 AGTGGCTGTGGGAAGACACAACTGCTGATGTCAGAGGCTCTATGACCCACAGAG 1344
Qy GlyPheIleMetValaspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
Db GGGATGCTCAGTGTGATGACAGAGATATTAGGACCAATAAATGTAAGGTTTCTACGGGAA 1404
Qy HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283
Db ATCATTTGGTGTGGTACGAGCAACCTGTTATGTTCCCAACCATAGTAGCTGAACACATT 1464
Qy LysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303
Db CGCTATGGCCGTGAAATGTCAACCATGATGAGATTGAAAGCTGTCAAGGAAGCCCAT 1524
Qy AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323
Db GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCCTGGTTGGAGAGAGGG 1584
Qy AlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343
Db GCCCAGTTGAGTGGTGGGCAAGACAGAGGATCGCCATTGCACGTGCCCTGGTTGCGAAC 1644
Qy ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
Db CCCAGATCTCTCTGCTGGATGAGCCACGCTCAGCCITGGACACAGAAACGAGCGAGTG 1704
Qy ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383
Db GTTCAGGTGGCTCTGGATAAGCCAGAAAGGTCGGACCAACCATTTGTGATAGCTCATCGT 1764
Qy LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403
Db TTGCTACAGTTCGTATGCTGAGTGCATCCCTGTTTCGATGATGAGTCAITTTGGGAG 1824
Qy LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423
Db AAGGAATCATGATGAACCTAGAAAGAGAAAGCGCATTTACTTTCAAACCTGTGCACAA 1884
Qy Gln-----AspIleLysLysAlaAspGluGlnMetGluSerMetThr 437
Db CACAGCAGGAAATGAAGTTGAATTAGAAATGCAGCTGATGAATCCAAAGTGAAT 1944
Qy TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal----- 452
Db GATGCCCTTGGAAATGCTTCAATGATGATTCAAGATCCAGTCTAATAAGAAAAAGCACT 2004
Qy LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys 470
Db CGTAGGAGTGTCCGTGGATCACAGCCCAAGAGAGAAAGCTTAGTACCAAGAGGCTCTG 2064
Qy GluIleSerLeuProGluValSerLeuLysIleLeuLysLeuAsnLysProGluTyr 490
Db GATGAAGTATATCTCAGTTTCTTTTGAGGATATGAGCTTAATTTAACTGAATGG 2124
Qy ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 510
Db CCTATTTTGTGTGTGTTTGTGCTATTAATGATTAATAATGAGCGCTGCAACAGCATTT 2184
Qy SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrLeu 529
Db GCAATATATTTCAAAGATATATAGGGGTTTTTACAGAAATGATGATCTCGAACAAAA 2244
Qy LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 549
Db CGACAGATATGATTAATCTTTTCACTATTGTTTCTACCCCTTGGAAATTTTCTTTTAT 2304
Qy SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 569
Db ACATTTTCTCAAGGTTTCACATTTTGGCAAGCTGGAGAGATCTCTCAACAGCGGCTC 2364
Qy ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 589
Db CGATACATGTTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTGTATGACCTAAA 2424

Qy AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 609
Db AACACCACTGGAGCATTTACTACAGGCTCGCCAAATGATGCTCTCAAGTTAAAGGGCT 2484
Qy ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629
Db ATAGGTTCCAGGCTCTCTGTAATTTACCCAGAAATATAGCAAACTTTGGGACAGAAAT 2544
Qy IleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeu 649
Db ATATCTTTCATCTAUGTTGGCAACTAAACACTTTACTCTTAGCAATTTGACCCATCAT 2604
Qy AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 669
Db GCAATAGCAGGAGTTTCTGAAATGAAATGTTGCTGGACAAGCACTGAAAGATAAGAAA 2664
Qy GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 689
Db GAACTAGAGGCTGCGGAAGATCGCTACTGAAGCAATAGAAAACCTTCCGAACCGTTGT 2724
Qy SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 709
Db TCTTTGACTCAGGAGCAGAACTTTGAAACATATGATGCTCAGAGTTTGCAGGTACCATAC 2784
Qy ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 729
Db AGAACTCTTTGAGGAAGACACACATCTTTTGAATTTACATTTTCTTCCACCCAGGCAATG 2844
Qy IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrIleuIleGlnAlaGlyArg 749
Db ATGTATTTTCTATCTGCTGATGTTTCCGTTTGGAGCCTACTTGTGGGCACATAAATC 2904
Qy MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 769
Db ATGAGTTTGTAGATGTTCTGTGATGATTTTTCAGCTGTTGTCTTTGTGCTGCGCGTG 2964
Qy GlyIleThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaHisLeu 789
Db GGGCAAGTCAGTTTCATTTGCTCTGCTACTATGCCAAAGCCAAATATATCAGCAGCCACATC 3024
Qy PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 809
Db ATCATGATCATTTGAAAAAACCCCTTTGATTGACAGCTACAGCAGCGAAGGCTTAATGCCG 3084
Qy AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 829
Db AACACATTGGAAGAAATGTACATTTGTTGTAAGTTGTATTTCAACTATCCACCCGACCG 3144
Qy AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 849
Db GACATCCAGTGTCTCAGGAGTGTGAGCTGAGCTGAGAGTGAAGAGGCCAGAGCTGCTCTG 3204
Qy ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 869
Db GTGGCAGCAGTGGCTGTGGGAAGACACAGTGGTCCAGCTCTCTGGAGCGGTTTCTACGAC 3264
Qy ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 889
Db CCCTTGGCAGGAGAAATGCTGCTGTGATGCAAGAAATAAAGCGACTGAATGTTCTAGTGG 3324
Qy LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909
Db CTTCCGAGCACCTGGGCATCGTGTCCAGAGGCCCATCTCTGTTGACTGCAGCATTTGCT 3384
Qy GluAsnIleAlaTyrGlyAspAsnSerArgValProLeuAspGluIleLysGluAla 929
Db GAGAACATTGCTTATGCAGACAAACAGCCGGTGGTGTCTACAGGAAGAGATCGTGAGGCA 3444
Qy AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 949
Db GCAAGAGGAGCCCAACATACATGCTTTCATCGATCTACGCTCTGCTAAATAATATACGCTAAA 3504


```
Qy 950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaAlaArg 969
Db 3505 GTAGGACAAAGCACTCAGCTCTGTGGCCAGAAACAAAGCAATGCCATAGCTCGT 3564
Qy 970 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989
Db 3565 GCCCTGTGACACGCTCATATTTGCTTTGGATGAAGCCAGCTCAGCTCTGGATACA 3624
Qy 990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009
Db 3625 GAAAGTGAAGAGGTGTCCAGAAAGCCCTGGACAAGCCAGAGAGCCGACCTGCRAT 3684
Qy 1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1029
Db 3685 GTGATTGCTCACCCGCTGTCCACCATCCAGATGACAGCTTAATAGTGGTTCAGAA 3744
Qy 1030 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1049
Db 3745 GGCAGAGTCAAGGAGCATGGCAGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTT 3804
Qy 1050 LysLeuValAsnAlaGlnSer 1056
Db 3805 TCAATGGTCAGTGTCCAGGCT 3825
RESULT 11
US-09-306-417-1
; Sequence 1, Application US/09306417
; Patent No. US20020103144A1
; GENERAL INFORMATION:
; APPLICANT: Heinrich-Pette-Institut
; TITLE OF INVENTION: Retroviral Gene Transfer Vectors
; FILE REFERENCE: P50491
; CURRENT APPLICATION NUMBER: US/09/306,417
; CURRENT FILING DATE: 1999-05-06
; EARLIER APPLICATION NUMBER: DE 198 22 115
; EARLIER FILING DATE: 1998-05-08
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 8630
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: proviral
; OTHER INFORMATION: plasmid DNA
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(160)
; OTHER INFORMATION: plasmid backbone (pUC)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (161)..(677)
; OTHER INFORMATION: 5'-LTR
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (532)..(1219)
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (1220)..(5062)
; OTHER INFORMATION: m4 mdr-1 cDNA
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (5215)..(5774)
; OTHER INFORMATION: 3'-LTR
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (5775)..(8630)
; OTHER INFORMATION: plasmid backbone (pUC)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(8630)
; OTHER INFORMATION: retroviral expression vector Sfbeta71m4
US-09-306-417-1
```

```
Alignment Scores:
Pred. No.: 5,71e-280 Length: 8630
Score: 3100.50 Matches: 595
Percent Similarity: 75.54% Conservative: 211
Best Local Similarity: 55.76% Mismatches: 218
Query Match: 58.17% Indels: 43
DB: 10 Gaps: 6
US-09-873-409-4 (1-1058) x US-09-306-417-1 (1-8630)
Qy 4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluVal 23
Db 1931 TCATTTACTGATAAAGAACTCTTAGCGGTATGCAAAAGCTGGACAGTAGCTGAAGAGTC 1990
Qy 24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuArgSer 43
Db 1991 TTGGCAGCAATTAAGAATGTGATTGTCATTTGGAGGACAAAGAAAGAACTTGAA----- 2044
Qy 44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63
Db 2044 ----- 2044
Qy 64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysaspPheGlyIleLys 83
Db 2045 -----AGGTACAAACAAAAATTTAGAAAGAGCTAAAGAAATTTGGGATAAAG 2089
Qy 84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
Db 2090 AAGACTATTACAGCCAATATTTCTATAGGTGCTGCTTCTCGTGTATCTATGATCTTAT 2149
Qy 104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123
Db 2150 GCTCTGGCCTTCTGTATGGACACCTTGGTCTCTCAGGGAA-----TATCTATT 2203
Qy 124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
Db 2204 GGCAAGTACTCCTGATTCTTTCTGATTAAATTTGGGGCTTTTAGTGTGGACAGSCA 2263
Qy 144 ValProHisPheGluThrPheAlaIleAlaArgValAlaPheHisIlePheGlnVal 163
Db 2264 TCTCCAAGCATTTGAAGCATTTGCAAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 2323
Qy 164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
Db 2324 ATTGATAATTAAGCCAGTATTGACAGCTATTCGAGAGTGGGCACAAACCCAGATAATT 2383
Qy 184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
Db 2384 AAGGGAAATTTGGAAATTCAGAAATGTTCACTTCAGTTACCCTATCTCGAAAGAAAGTTAAG 2443
Qy 204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
Db 2444 ATCTTGAAGGGCCTGAACCTGAAGGTGACAGTGGGACAGCGTGGCCCTGGTGGAAAC 2503
Qy 224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243
Db 2504 AGTGGCTGTGGAGAGACACACAGCTCAGCTGATGAGAGGGCTCTATGACCCACAGAG 2563
Qy 244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
Db 2564 GGGATGCTCAGTGTGTGATGACAGGATATTAGGACCACTAAATGTAAAGTTTCTACGGGAA 2623
Qy 264 HisIleGlyValValSerGlnGluProValIleuPheGlyThrThrIleSerAsnIle 283
Db 2624 ATCATTTGGTGTGGTGTGAGTACAGAACCTGTATTGTTTCCACACGATAGCTGAAACAT 2683
Qy 284 LysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303
Db 2684 CGCTATGCCGTGAAATGTCCACCATCGATGAGATTCAGAAAGCTGTCAAGGAAGCAAT 2743
Qy 304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323
Db 304 ----- 323
```


Db 2744 GCCTATGACCTTTATCATGAACCTGCCTCAATAAATTTGACACCCCTGGTTGAGACAGAGGG 2803
Qy 324 AlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaAlaLeuValArgAsn 343
Db 2804 GCCCAGTTGAGTGGTGGCAGAAAGCAGAGATCGCCATTGCAGTGCCTGGTTCCGAAC 2863
Qy 344 ProIysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
Db 2864 CCCAAGATCCTCGTGGATGAGGCACCTCAGCCTTGACACAGAAAGCGAAGCAGTG 2923
Qy 364 ValGlnAlaAlaLeuGlnLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383
Db 2924 GTTCAGTGCTCTGGATAGGCAGAAAGTCGGACCACTTGTGATGATCTATCGT 2983
Qy 384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403
Db 2984 TTGCTACAGTTGTAATGCTGACGTCATCGCTGTTTCGATGATGAGTGTGGAG 3043
Qy 404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrSerLeuValMetSer 423
Db 3044 AAAGGAAATCATGATGAACCTCATGAAGAGAAAGGCACTTACTTCAAACTTGTCAAAATG 3103
Qy 424 Gln-----AspIleLysAlaAspGluGlnMetGluSerMetThr 437
Db 3104 CAGACAGCAGAAATGAAGTTGAATTAGAAATGACGCTGATGAATCCAAAGTGAAT 3163
Qy 438 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal----- 452
Db 3164 GATGCCTTGGAAATGCTTCAATGATTCAGATCCAGTCTAATAGAAAAGATCAACT 3223
Qy 453 ---LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys 470
Db 3224 CGTAGGAGTGTCCGTGATCACAGGCCAACAGACAGAAAGCTTAGTACCAAGAGGCTCTG 3283
Qy 471 GluIleSerLeuProGluValSerLeuLysIleLeuLysLeuAsnLysProGluTrp 490
Db 3284 GATGAAGTATACCTCCAGTTTCTTTGGAGGATATGAAGCTAAATTTAACTGAATGG 3343
Qy 491 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 510
Db 3344 CCTATTGTTGTTGGTGTATTTTGGCCATTATATAATGGAGGCTGCAACCGCATTT 3403
Qy 511 SerIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrLeu 529
Db 3404 GCAATAATATTTCAAAGATTATAGGGTGTTTACAAAGATTGATGATCCTGAAACAAA 3463
Qy 530 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 549
Db 3464 CGACAGAAATAGTAACCTGTTTTCATATTGTTTCTAGCCCTTGGAAATATTCTTTATT 3523
Qy 550 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 569
Db 3524 ACATTTTCTTCAAGTTTCACATTTGGCAAGCTGGAGAGATCTCACCAAGCGGCTC 3583
Qy 570 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 589
Db 3584 CGATACATGTTTCCGATCCATGCTCAGACAGGATGTGATGTTGTTGATGCCCTAAA 3643
Qy 590 AsnSerThrGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 609
Db 3644 AACACCACTGGACATTTGACTTACAGCTGCCAATGATGCTCCTCAAGTTTAAAGGGGCT 3703
Qy 610 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629
Db 3704 ATAGTTTCCAGGCTTGCTGTAATTTACCAGATATAGCAATCTTGGCAGAGATTAAT 3763
Qy 630 IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 649
Db 3764 ATATCTTTCATGTTGTCACACTTAACACTGTTACTTCTTAGCAATTGATCCCATCAT 3823
Qy 650 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 669
Db 3824 GCAATAGCAGGAGTTGTTGAAATGAAATGTTGTTCTGGCAAGCACTGAAAGATAAGAAA 3883

Qy 670 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 689
Db 3884 GAACTAGAAGAGTGTCTGGGAAGATCGCTACTGAAGCAATAGAAAACCTTCCGAACCGTTGT 3943
Qy 690 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 709
Db 3944 TCTTTGACTCAGAGAGCAAGATTTGAACATATGATGTCTCAGAGTTTGCAGGTACCATATC 4003
Qy 710 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 729
Db 4004 AGAACTCTTTGAGGAAGACACACATCTTTGGAAATACATTTTCTTCCACCGGCAATG 4063
Qy 730 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyValTyrIleLeuIleGlnAlaGlyArg 749
Db 4064 ATGTATTTTCTCTGCTGATGTTTCCGTTTGGAGCCTACTTGTGTGGCACATAAATC 4123
Qy 750 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 769
Db 4124 ATGAGCTTTGAGGATGTTCTGTAGTATTTTTCAGCTGTTGTCTTGTGTCATGGCCGTG 4183
Qy 770 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaHisLeu 789
Db 4184 GGGCAAGTCAGTTCATTTGCTCTGACTATGCCAAAGCCAAATATATCAGCAGCCACATC 4243
Qy 790 PheAlaLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 809
Db 4244 ATCATGATCATTTGAAAAACCCCTTTGATTTGACAGCTACAGCACGGAAGCCCTAATGCCG 4303
Qy 810 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 829
Db 4304 AACACATTGGAAGAAATGTACATTTGTGTAAGTTGTATTCAACTATCCACCCGACCG 4363
Qy 830 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 849
Db 4364 GACATCCAGTGTCTCAGGACTGAGCTGGAGGTGAAGAGGCCAGAGCTGGCTCTG 4423
Qy 850 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 869
Db 4424 GTGGCAGCAGTGGCTGTGGGAAGACACAGTGGTCCAGCTCTCGAGCGGTTCTACGAC 4483
Qy 870 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 889
Db 4484 CCTTGGCAGGGAATGCTGTGTTGATGCAAAAGAAATTAAGACGACTGAATGTTCAGTGG 4543
Qy 890 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909
Db 4544 CTCGAGACACCTGGGCATCTGTTCCAGAGCCCATCTCTGTTTACTGCACATTTGCT 4603
Qy 910 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 929
Db 4604 GAGAACATTCCTATCGAGACACACCGCGGTGTCTCACAGGAAGAGATCGTGAGGCA 4663
Qy 930 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 949
Db 4664 GCAAAGGAGGCCAACATACATGCTTTCATCGAGTCACTGCTTAATAATATAGACATAA 4723
Qy 950 ValGlyLeuLysGlyValaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 969
Db 4724 GTAGGAGACAAAGGAATCAGCTCTCTGTGTGGCCAAACACAGCATTTGCCATGCTCGT 4783
Qy 970 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989
Db 4784 GCCCTTTGAGACAGCTCATATTTTGTCTTTGATGAAGCCAGCTCAGCTCTGGATACA 4843
Qy 990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009
Db 4844 GAAAGTGAAGAGTTGTCCAAAGAGGCCCTCGACAAAGCCAGAGAGCGCCACCTGCAT 4903
Qy 1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1029
Db 4904 GTGATTTGCTCACCGCTTCCACCATCCAGATGCAGACTTAATAGTGGTGTTCAGAAT 4963

Db 2984 TTGCTACAGTTTCGTAAATGCTGAGTCATCGCTGGTTTCGATGATGAGTCATTTGGAG 3043
Qy 404 LygGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423
Db 3044 AAAGGAAATCATGATGAATCATGAAAGGAAAGGCAATTTACTTCAAACTTTGTCACAATG 3103
Qy 424 Gln-----AspIleLysLysAlaAspGluGlnMetGluSerMetThr 437
Db 3104 CAGACAGCAGGAAATGAAGTTGAATTAGAAATGCGAGCTGATGAATCCAAAGTGAAT 3163
Qy 438 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal----- 452
Db 3164 GATGCTTGGAAATGCTTCAATGATTCAGATCCAGTCTAATAAGAAAAAGATCAACT 3223
Qy 453 ---LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys 470
Db 3224 CGTAGGAGTTCGCTGGATCACAAGCCCAAGCAGACAGAAAGCTTAGTACCAAGAGGCTCTG 3283
Qy 471 GluIleSerLeuProGluValSerLeuLysIleLeuLysLeuAnLysProGluTyr 490
Db 3284 GATGAAAGTATACCTCCAGTTTCTTTTGGAGGATTAAGAAGCTAAATTTAACTGAATGG 3343
Qy 491 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 510
Db 3344 CCTATTGTTGTTGGTGATTTTGTGGCATTATTAATGGAGCCCTGCACAGCAATTT 3403
Qy 511 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 529
Db 3404 GCAATAATATTTCAAAGATTATAGGGGTTTACAGAAATTCATGATCCTGAAACAAAA 3463
Qy 530 LysHisAspAlaGluLysSerMetIlePheValIleLeuGlyValIleCysPheVal 549
Db 3464 CGACAGAAATAGTAACCTGTTTCACTATTGTTCTAGCCCTTGGAAATTTCTTTTATT 3523
Qy 550 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 569
Db 3524 ACATTTTCTTCAAGGTTTCACATTTGGCAAGCTGGAGAGATCTTCACCAAGCGGCTC 3583
Qy 570 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGluLysGlu 589
Db 3584 CGATACATGGTTTTCGATCCATGCTCAGACAGGATGAGTTGGTTTGATGCCCTAAA 3643
Qy 590 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 609
Db 3644 AACACCACTGGAGCATTTGACTACAGGCTCGCCAATGATGCTCAAGTTAAAGGGGCT 3703
Qy 610 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629
Db 3704 ATAGGTTCCAGGCTTCTGTATATTTACCAGAAATATAGCAATCTTGGGACAGAAATATT 3763
Qy 630 IleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeu 649
Db 3764 ATATCTTTCATATGTTGGCAACTAACACTGTTACTCTTAGCAATTTGATCCCATCAT 3823
Qy 650 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 669
Db 3824 GCAATAGCAGGAGTTGTGAAATGAAATGTTGCTCGCAGCAGCACTGAAAGATAGAAAA 3883
Qy 670 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 689
Db 3884 GAACTAGAAGGTGCTGGGAAGATCGCTACTGAAAGCAATAGAAAACTTCCGAACCGTTGTT 3943
Qy 690 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 709
Db 3944 TCATTGACTCAGGAGCAGAGTTTGAACATATGATGCTCAGAGTTTGCAGGTACCATAC 4003
Qy 710 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 729
Db 4004 AGAAACTCTTTGAGAAAGCACACATCTTTTGAATTAACATTTTCTTCCACCAGGCAATG 4063
Qy 730 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 749
Db 4064 ATGTATTTTTCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTTGGTGCCACATAAACTC 4123

Qy 750 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 769
Db 4124 ATGAGCTTTGAGATGTTCTGTAGTATTTCAGCTGTTGTCTTGTGGTCATGGCGGTG 4183
Qy 770 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 789
Db 4184 GGCAGATCAGTTCATTTGCTCTGCTGATATGCCAAGCCAAAATATCAGCAGCCCATC 4243
Qy 790 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 809
Db 4244 ATCATGATCATTTGAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGCCCTAATCCG 4303
Qy 810 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 829
Db 4304 AACACATTGGAAGAAATGTACATTTGGTGAAGTTGTATTCAACTATCCACCCGACCG 4363
Qy 830 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 849
Db 4364 GATCCTCCAGTCTTTCAGGACTGAGCTGGAGTGAAGAAAGGCCAGACGCTGGCTCTG 4423
Qy 850 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 869
Db 4424 GTGGCAGCAGTGGCTGTGGGAAGACACAGTGTCCAGCTCTCGAGCGGTTCTACGAC 4483
Qy 870 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTyr 889
Db 4484 CCCTTGGCAGGAAAGTGTCTGTGTCGAAAGATAAAGCGACTGAAATGTTTCAGTGG 4543
Qy 890 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909
Db 4544 CTCGAGACACATCGGCGATCGTGTCCCAGAGCCCATCTCTGTGACTGCAGCATTTGT 4603
Qy 910 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 929
Db 4604 GAGAAATTGGCTATGGAGACACACCGGGTGTGTACAGGAGAGATCGTGAGGCGCA 4663
Qy 930 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 949
Db 4664 GCAAGGAGGCCAACATACATGCTTTCATCGACTCATGCTAATAAATATAGCACTAAA 4723
Qy 950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 969
Db 4724 GTAGGAGACAAAGAACTCAGCTCTCTGTGGCCAGAACCAACGCAATTCCTAGCTGT 4783
Qy 970 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989
Db 4784 GCCTTGTAGACAGCTCATATTTTGTGATGAAGCCACGCTCAGCTCTGGATACA 4843
Qy 990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009
Db 4844 GAAAGTGAAAGGTTGTCCAAGAAAGCCCTGGACAAAGCCAGAGAACGCCGCTGCATT 4903
Qy 1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValLeuHisAsn 1029
Db 4904 GTGATTGCTCACCGCTGTCCACATCCAGATTCAGACTTAATAGTGGTGTTCAGAAAT 4963
Qy 1030 GlyLysIleLysGluGlnGlyThrHisGlnLeuLeuArgAsnArgAspIleTyrPhe 1049
Db 4964 GGCAAGTCAAGGAGCATGGCAGCATCAGCAGCTGTGCGACAGAAAGGCACTCTATTTT 5023
Qy 1050 LysLeuValAsnAlaGlnSer 1056
Db 5024 TCAATGGTCACTGCCAGGCT 5044

RESULT 13

US-09-866-866A-7

; Sequence 7, Application US/098666866A

; Patent No. US2002010224A1

; GENERAL INFORMATION:

; APPLICANT: Sorrentino, Brian

; APPLICANT: Schuetz, John

; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells

FILE REFERENCE: 1340-1-021CIP2
 CURRENT APPLICATION NUMBER: US/09/866,866A
 CURRENT FILING DATE: 2001-08-30
 PRIOR APPLICATION NUMBER: 09/584,586
 PRIOR FILING DATE: 2000-05-31
 PRIOR APPLICATION NUMBER: PCT/US99/11825
 PRIOR FILING DATE: 1999-05-27
 PRIOR APPLICATION NUMBER: 60/086,988
 PRIOR FILING DATE: 1998-05-28
 NUMBER OF SEQ ID NOS: 27
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO 7
 LENGTH: 4788
 TYPE: DNA
 ORGANISM: Mus musculus
 US-09-866-866A-7

Alignment Scores:
 Pred. No.: 2,06e-279 Length: 4788
 Score: 3090.50 Matches: 588
 Percent Similarity: 75.12% Conservative: 215
 Best Local Similarity: 55.00% Mismatches: 219
 Query Match: 57.98% Indels: 47
 DB: 10 Gaps: 6

US-09-873-409-4 (1-1058) x US-09-866-866A-7 (1-4788)

Qy 4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluVal 23
 Db 700 TCAATTTACTGATAAGAACTCCATGCTTATGCAAAAGCTGGAGCAGTTGCTGAAGAAGTC 759
 Qy 24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43
 Db 760 TTACGAGCCATCAGAACTGTGATTCGTTGGAGACAAAAGAGAACTTGAA----- 813
 Qy 44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63
 Db 813 ----- 813
 Qy 64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
 Db 814 -----AGGTACAATAACAACTTGGAGAAAGCTGAAAGGCTGGGGATAAAG 858
 Qy 84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
 Db 859 AAAGCTATCAGCGCCAACTCTCCATGGGTGCGAGCTTCTCCTTATCATGTCATCAT 918
 Qy 104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123
 Db 919 GCTCTGGCATTCGTATGGGACTTCCTTGGTCACTCCAAAGAA-----TACTCTATT 972
 Qy 124 GlyThrValLeuAlaValPhePheSerValIleHisSerTyrCysIleGlyAlaAla 143
 Db 973 GGACAAAGTGCTCACTGTCTCTTTCCGTGTTAATGGAGCATTCAGTGTGGACAGCA 1032
 Qy 144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163
 Db 1033 TCTCCAAATATTGAAGCCTTCGCCAATGACGAGGAGCAGCTTATGAAGTCTTCAAAATA 1092
 Qy 164 IleAspLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
 Db 1093 ATTGATAATAGCCCATATAGACAGCTTCTCAAGAGTGGGCACAAACGACACATA 1152
 Qy 184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
 Db 1153 CAAGGAATCTCGAATTTAAGAATATTCACCTTCAGTTACCCATCTCGAAAGAGTTCAG 1212
 Qy 204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
 Db 1213 ATCTTGAAGGGCCTCAATCTGAAGGTGAAGCGGACGACGCGTGGCCCTGTTGGCAAC 1272
 Qy 224 AsnGlySerGlyLysSerThrValGlnLeuLeuGlnArgLeuTyrAspProAsp 243
 Db -----

Db 1273 AGTGGCTGTGAAAAAGACAACTGTCCAGCTGATGCAAAAGGCTCTACAGCCCCCTAGAT 1332
 Qy 244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
 Db 1333 GGCATGTCAGTATCGACGACAGGACATCAGAACCATCAATGTGAGGTATCTCAGGGAG 1392
 Qy 264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnIle 283
 Db 1393 ATCATTGGTGTGGTAGTCAGGAACCTGTCTGTTGCCACACGATCGCGGAGAACATT 1452
 Qy 284 LysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303
 Db 1453 CGCTATGCCGAGAGATGTCCACATGGATGAGATGAGAAAGCTGTCAAGGAAGCCAA 1512
 Qy 304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323
 Db 1513 GCCTATGACTTCATCAAACTGCCCCACCAATTTGACACCTTGGTGGTGAGAGGG 1572
 Qy 324 AlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343
 Db 1573 GCGCACGTGAGTGGGACAGAAACAGAGAAATGCCATTGCGCGGCCCTGGTCCGCAAT 1632
 Qy 344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
 Db 1633 CCCAAGATCCTTTTGTGGACGAGGCCCTCAGCCCTGGATACAGAAAGTGAAGCTGTG 1692
 Qy 364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383
 Db 1693 GTTCAGCGCCGACATGAGTAAGGCTAGAGAAAGCCGACACCATTTGATGATGATCATCGC 1752
 Qy 384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403
 Db 1753 TTGTCTACCGTTCGTAATGCTGACGTCATCTGCTGTTTGTGTTGCTGTTGTTGAG 1812
 Qy 404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423
 Db 1813 CAAGGAATCATGATGATGCTCATGAGAGAAAGGCAATTTACTTCAAACTTGTGATGACA 1872
 Qy 424 Gln-----AspIleLysLysAlaAspGluGlnMet 433
 Db 1873 CAGACAGCAGGAATGAATTTGAATTTAGGAATAGAGCTTGTAAATCTAAGGATGAAT 1932
 Qy 434 GluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerValLys 453
 Db 1933 GATAATTTAGACATGCTCTTCAAAAGATTCCAGATCCAGTCTAATAAGAAGA-----AGA 1986
 Qy 454 SerLysSerAspPheIle-----AspLysAlaGluSerThrGln 468
 Db 1987 TCAACTCGCAAAAGCATCTGTGGACCATATGACCAAGACAGGAGCTTAGTACCAAGAG 2046
 Qy 469 SerLysGluIleSerLeuProGluValSerLeuLysIleLeuLysLeuAsnLysPro 488
 Db 2047 GCCTGATGAAGATGATACCTCCAGTTCCTTTGGCGGATCCTTGAAGTTGAATCAACT 2106
 Qy 489 GluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPro 508
 Db 2107 GAATGGCTTATTTTGGTTGGTATATCTGTGCCATAATAATGGAGGCTTACAGCA 2166
 Qy 509 ValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThr 527
 Db 2167 GCATTCCTCGTAATATTTTCAAAAGTTGTAGGGGTTTTTACAAATGGTGGCCCCCTGAA 2226
 Qy 528 ThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCys 547
 Db 2227 ACCCAGCGCAGACAGCAACTGTTTCTGTGTGTCTTGTCTTGTCTTGTGGATCATTTCT 2286
 Qy 548 PheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMet 567
 Db 2287 TTCATTACATTTTCTTCAGGCTTCACATTTGCAAGGCTGGAGAGATCCTCACCAG 2346
 Qy 568 ArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGlu 587
 Db 2347 CGACTCCGATACATGGTTTTCAAATCCATGCTGAGACAGGATGTGAGTGTGTTGATGAC 2406

Qy	948	ThrcInValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIle	967
Db	3487	ACCAGATAGGCACAAAGCCACTCAGCTGTGGTGGCGAAGCAGCATGCCCATC	3546
Qy	968	AlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeu	987
Db	3547	GCAGCGCCCTCGTCAGACAGCTCACATTTTACTTCTTGACGAAGCAACATCAGCTCTG	3606
Qy	988	AspAsnASerGlnLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThr	1007
Db	3607	GATACAGAAGAGTGAAAGGTTGTCAGGAAGCGCTGCGACAAGCCAGGCAAGCGGCACC	3666
Qy	1008	CysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAapLeuIleValValLeu	1027
Db	3667	TGCATTGTGATGCTCACCGCCCTGTCCACATCCAGAACCGGAGCTTGATCGTGGTGATT	3726
Qy	1028	HisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspile	1047
Db	3727	CAGAACGGCAAGGTCAAGGAGCAAGCCACCCACAGCAGCTGCTGGCGCAGAGGGCATC	3786
Qy	1048	TyrPhelLysLeuValAsnAlaGlnSer	1056
Db	3787	TACTTCTCAATGGTCAGTGTGCAGGCT	3813
RESULT 14			
US-09-880-107-2299			
; Sequence 2299, Application US/09880107			
; Patent No. US20020142981A1			
; GENERAL INFORMATION:			
; APPLICANT: Horne, Darci T.			
; APPLICANT: Vockley, Joseph G.			
; APPLICANT: Scherf, Uwe			
; APPLICANT: Gene Logic, Inc.			
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer			
; FILE REFERENCE: 44921-5028-WO			
; CURRENT APPLICATION NUMBER: US/09/880,107			
; CURRENT FILING DATE: 2001-06-14			
; PRIOR APPLICATION NUMBER: US 60/211,379			
; PRIOR FILING DATE: 2000-06-14			
; PRIOR APPLICATION NUMBER: US 60/237,054			
; PRIOR FILING DATE: 2000-10-02			
; NUMBER OF SEQ ID NOS: 3950			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 2299			
; LENGTH: 3924			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 M23234			
US-09-880-107-2299			
Alignment Scores:			
Pred. No.:	5,02e-279	Length:	3924
Score:	3085.00	Matches:	597
Percent Similarity:	74.72%	Conservative:	198
Best Local Similarity:	56.11%	Mismatches:	229
Query Match:	57.88%	Indels:	40
DB:	10	Gaps:	5
US-09-873-409-4 (1-1058) x US-09-880-107-2299 (1-3924)			
Qy	4	SerLeuThrSerLysGlnLeuSerAlaTyrrSerLysAlaGlyAlaValAlaGluVal	23
Db	750	GCATTTAGTGACAAAAGTAAGTAGCTGCTTTATGCAAAGCAGGCGCGCTGGCAGAAGGCT	809
Qy	24	LeuSerSeriLeArghrValIleAlaPheArGAlaGlnGluLysGluLeuGlnArgSer	43
Db	810	CTGGGGCCATCAGGACTGTGATAGCTTTCCGGGGCCAGAACAAAGAGCTGAA-----	863
Qy	44	PhelLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys	63
Db	863	-----	863

QY 64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
Db 864 -----AGGTATCAGAAACATTTAGAAAATGCCAAAGAGATTGGAAATAAA 908
QY 84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
Db 909 AAAGCTATTTCAGCAAAATTTCCATGGGTATTCCCTTCTGTTTAATATATATGCATCATAT 968
QY 104 GlyLeuAlaPheTyrThrGlyThrSerLeuIleLeuAsnGlyGluProGlyThrIle 123
Db 969 GCATCGCCCTTCGTGTATGGATCCACTCTAGTCATATCAAAAGAA-----TATACATATT 1022
QY 124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
Db 1023 GGAATGCAATGACAGTTTTTTTTTCAATCCTAATTGGAGCTTTCAGTGTTGGCAGGCT 1082
QY 144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163
Db 1083 GCCCATGTATTGATGCTTTTGGCAATGCAAGAGGAGCAGCATATGTGATCTTTGATATT 1142
QY 164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
Db 1143 ATTCATTAATATCCTAAATATGACAGTTTTTTCAGAGAGAGACAAACACGACGATC 1202
QY 184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
Db 1203 AAAGGGAATTTGGAGTTCAATGATGTTTCACTTTTCTTACCCTTCTCGAGCTAACGTCAG 1262
QY 204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
Db 1263 ATCTTTGAAGGCGCTCAACCTGAAGGTGCAGAGTGGCGAGCGTGGCCCTGGTTGGAGT 1322
QY 224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243
Db 1323 AGTGGCTGTGGGAAGAGACAAACGGTCCAGCTGATACAGAGCTCTATGACCCCTGATGAG 1382
QY 244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
Db 1383 GGCACAATTAACTTGTATGGCCAGGATATTAGGAACTTTAAATGTAACATCTCTGAGGAA 1442
QY 264 HisIleGlyValValSerGlnProValLeuPheGlyThrIleSerAsnAsnIle 283
Db 1443 ATCATTTGTTGGTGTGAGTCAGGCGGTGCTGTTTCCACCAATTTGCTGAAATATT 1502
QY 284 LysTyrGlyArgAspAspValTyrAspGluMetGluArgAlaAlaArgGluAlaAsn 303
Db 1503 TGTATTGGCCGTGGAATGTAACTGATGAGATAAAGAGCTGTCAAAAGAGGCCAAC 1562
QY 304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyLysGly 323
Db 1563 GCCTATGAGTTTATCATGAATTAACACAGAAATTTGACACCCCTGGTTGGAGAGAGGG 1622
QY 324 AlaGlnMetSerGlyGlyGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343
Db 1623 GCCCAGCTGAGTGTGGGAGAGCAGAGGATCGCCATTGCACGTGCGCTGGTTCGCAAC 1682
QY 344 ProlLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
Db 1683 CCCAAGATCCCTTCTGCTGGATGAGGCCACGTCAGATTGGACACAGAAAGTGAAGCTGAG 1742
QY 364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383
Db 1743 GTACAGGAGCTCTGGATTAAGGCCAGAGAGCGCCGACCACTTGTGATAGCACACCGA 1802
QY 384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403
Db 1803 CTGTCTACGGTCCGAAATGTCAGATGTCATCGCTGGGTTTGGAGTGGAGTAATTGTGGAG 1862
QY 404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuVal----- 421
Db 1863 CAAGGAAGCCACAGCGAACTGATGAAGAAAGGAGGGGTACTTCAAACTTTGTCAACATG 1922
QY 422 -----MetSerGlnAspIleLysLysAlaAspGluGlnMetGlu 434

Db 1923 CAGACATCAGGAAGCCAGATCCAGTCAGAAAGAAATTTGAACTAATGATGAAAGGTCGCC 1982
QY 435 SerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSer 454
Db 1983 ACTAGAAATGCCCCCAATAGGCTGGAATATCTCGCTATTATTAGGCAATCTACTCAGAAAAAC 2042
QY 455 IleLys---SerAspPheIleAspLysAla---GluGluSerThrGlnSerLysGluIle 472
Db 2043 CTTAAAAATTCACAAATGTCTCAGAAGAGCCTTGATGTGGAACCGATGGATTTGAAGCA 2102
QY 473 SerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTyrProPhe 492
Db 2103 AATGTGCCACAGTGTCTTCTCTGAAGGTCTCTGAAATCTGAAATGAAACAGAAATGCCCCATC 2162
QY 493 ValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIle 512
Db 2163 TTTGTCTGGGAACAGATATGTGCATTTGCCAATGGGGGGCTTCAGCCGGCATTTTCAGTC 2222
QY 513 IlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThrLeuLysHisAsp 532
Db 2223 ATATTCTCAGAGATCATAGCGATTTTGGACAGGCGATGATGCAGTGAAGCAGCAGAAG 2282
QY 533 AlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPhe 552
Db 2283 TGCAACATATTTCTTGTGATTTCTTATTTCTGGGAATATTCTTTTACTTTCTTCTTC 2342
QY 553 MetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeu 572
Db 2343 CTTTCAGGGTTTCAGTTTGGGAAGCTGGCGAGATCTCCACGAGAGCTGCGGTCAATG 2402
QY 573 AlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGluLysGluAsnSerThr 592
Db 2403 GCITTTAAAGCAATGCTTAAGCAGGACATGAGCTGGTTTGTATGACCATATAAAACAGTACT 2462
QY 593 GlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySer 612
Db 2463 GGTGCACCTTCTACAGACTTGGCCACAGATGTGCTCCCAAGTCCCAAGGAGGCCACAGAAC 2522
QY 613 ArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPhe 632
Db 2523 AGGTGGCTTTTAAATTTGCACAGAAATATAGCTAACCTTGGAACTGGTATTATCATATCATTT 2582
QY 633 IleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThr 652
Db 2583 ATCTACGGTTGGCAGTTAAACCTATTGCTATTAGCAGTGTGTCCATATTATGCTGTGCA 2642
QY 653 GlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLys 672
Db 2643 GGAATTTGAAATGAAATTTGCTGGCTGGAATGCCAAAAGAGATAAAAAAGAACTGGAA 2702
QY 673 HisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThr 692
Db 2703 GCTCTCGAAAGATTTGCAACAGAGGCAATAGAAAATATTAGGACAGTGTGTCTTTGACC 2762
QY 693 ArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThr 712
Db 2763 CAGAAAGAAAATTTGAATCAATGATGTTGAAAAATTTGATGGACCTTACAGGAATTTCT 2822
QY 713 SerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPhe 732
Db 2823 GTGCAGAAAGGCACACATCTATGGAATTTACTTTTAGTAGTCTCTCACAGCATTTTATGTATTT 2882
QY 733 AlaTyrAlaAlaGlyPheArgPheGlyValTyrIleLeuIleGlnAlaGlyArgMetThrPro 752
Db 2883 TCCATATGCCGGTGTGTTTTCGATTTGGTCATATCTCATTTGTAATGGACATATCGGCTTC 2942
QY 753 GluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThr 772
Db 2943 AGAGATGTTATCTGGTGTGTTTCTGCAATTTGTAATTTGGTGGAGTGGCTCTAGGACATGCC 3002
QY 773 LeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeu 792

3003	AGTTCATTCTCCAGACTATGCTAAAGCTAAGCTGTCTGCAGAGCCCACTTATTCAATGCTG	3062	3063	AGTTCATTCTCCAGACTATGCTAAAGCTAAGCTGTCTGCAGAGCCCACTTATTCAATGCTG	3062
Qy	LeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCys	812	3063	TTTGAAGAACAACCTCTGATTGACAGCTACAGCTGAAGAGGGGGCTGAAGCCCTGATAAAATTT	3122
Db	LeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCys	812	3063	TTTGAAGAACAACCTCTGATTGACAGCTACAGCTGAAGAGGGGGCTGAAGCCCTGATAAAATTT	3122
Qy	GluGlyAsnLeuGluPheArgGluValSerPheThrProCysArgProAspValPhe	832	813	GluGlyAsnLeuGluPheArgGluValSerPheThrProCysArgProAspValPhe	832
Db	GluGlyAsnLeuGluPheArgGluValSerPheThrProCysArgProAspValPhe	832	813	GluGlyAsnLeuGluPheArgGluValSerPheThrProCysArgProAspValPhe	832
Qy	GAAGGAAATATACTTAAATGAAGTCGTGTTCAACTATCCACCAGGCAACGTCGCCA	3182	3123	GAAGGAAATATACTTAAATGAAGTCGTGTTCAACTATCCACCAGGCAACGTCGCCA	3182
Db	GAAGGAAATATACTTAAATGAAGTCGTGTTCAACTATCCACCAGGCAACGTCGCCA	3182	3123	GAAGGAAATATACTTAAATGAAGTCGTGTTCAACTATCCACCAGGCAACGTCGCCA	3182
Qy	IleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySer	852	833	IleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySer	852
Db	IleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySer	852	833	IleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySer	852
Qy	GTGCTTCAGGGGCTGAGCTGGAGGTGAAGAAGGCCAGACACTAGCCCTGCTGGCGCAGC	3242	3183	GTGCTTCAGGGGCTGAGCTGGAGGTGAAGAAGGCCAGACACTAGCCCTGCTGGCGCAGC	3242
Db	GTGCTTCAGGGGCTGAGCTGGAGGTGAAGAAGGCCAGACACTAGCCCTGCTGGCGCAGC	3242	3183	GTGCTTCAGGGGCTGAGCTGGAGGTGAAGAAGGCCAGACACTAGCCCTGCTGGCGCAGC	3242
Qy	SerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyArgProValGln	872	853	SerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyArgProValGln	872
Db	SerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyArgProValGln	872	853	SerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyArgProValGln	872
Qy	AGTGGCTGTGGGAAGACACGGTGGTCTCAGCTCTCTGGAGCGGTCTTACGACCCCTCGCG	3302	3243	AGTGGCTGTGGGAAGACACGGTGGTCTCAGCTCTCTGGAGCGGTCTTACGACCCCTCGCG	3302
Db	AGTGGCTGTGGGAAGACACGGTGGTCTCAGCTCTCTGGAGCGGTCTTACGACCCCTCGCG	3302	3243	AGTGGCTGTGGGAAGACACGGTGGTCTCAGCTCTCTGGAGCGGTCTTACGACCCCTCGCG	3302
Qy	GlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSer	892	873	GlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSer	892
Db	GlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSer	892	873	GlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSer	892
Qy	GGACAGTGGCTTCTCATGGTCAAGAGCAAGAACAATCAATGTCCAGTGGCTCAGAGCT	3362	3303	GGACAGTGGCTTCTCATGGTCAAGAGCAAGAACAATCAATGTCCAGTGGCTCAGAGCT	3362
Db	GGACAGTGGCTTCTCATGGTCAAGAGCAAGAACAATCAATGTCCAGTGGCTCAGAGCT	3362	3303	GGACAGTGGCTTCTCATGGTCAAGAGCAAGAACAATCAATGTCCAGTGGCTCAGAGCT	3362
Qy	GlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIle	912	893	GlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIle	912
Db	GlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIle	912	893	GlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIle	912
Qy	CAACTCGGAATCGTGTCTCAGGAGCCTATCTCTATTGTGACTGCAGCATTTGCCGAGAAATAT	3422	3363	CAACTCGGAATCGTGTCTCAGGAGCCTATCTCTATTGTGACTGCAGCATTTGCCGAGAAATAT	3422
Db	CAACTCGGAATCGTGTCTCAGGAGCCTATCTCTATTGTGACTGCAGCATTTGCCGAGAAATAT	3422	3363	CAACTCGGAATCGTGTCTCAGGAGCCTATCTCTATTGTGACTGCAGCATTTGCCGAGAAATAT	3422
Qy	AlaTyrcGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAla	932	913	AlaTyrcGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAla	932
Db	AlaTyrcGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAla	932	913	AlaTyrcGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAla	932
Qy	GCCTATGGAGACAAACAGCCGGTGTGTATCACAGGATGAAATTTGTGAGTGCCAGCCAAAGCT	3482	3423	GCCTATGGAGACAAACAGCCGGTGTGTATCACAGGATGAAATTTGTGAGTGCCAGCCAAAGCT	3482
Db	GCCTATGGAGACAAACAGCCGGTGTGTATCACAGGATGAAATTTGTGAGTGCCAGCCAAAGCT	3482	3423	GCCTATGGAGACAAACAGCCGGTGTGTATCACAGGATGAAATTTGTGAGTGCCAGCCAAAGCT	3482
Qy	AlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyArgAsnThrGlnValGlyLeu	952	933	AlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyArgAsnThrGlnValGlyLeu	952
Db	AlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyArgAsnThrGlnValGlyLeu	952	933	AlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyArgAsnThrGlnValGlyLeu	952
Qy	GCCAAACATACATCTTTTCATCGAGAGCTTACCCCAACAATATGAAACAAGAGTGGAGAT	3542	3483	GCCAAACATACATCTTTTCATCGAGAGCTTACCCCAACAATATGAAACAAGAGTGGAGAT	3542
Db	GCCAAACATACATCTTTTCATCGAGAGCTTACCCCAACAATATGAAACAAGAGTGGAGAT	3542	3483	GCCAAACATACATCTTTTCATCGAGAGCTTACCCCAACAATATGAAACAAGAGTGGAGAT	3542
Qy	LysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeu	972	953	LysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeu	972
Db	LysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeu	972	953	LysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeu	972
Qy	AAGGGGACTCAGCTCTCAGGAGGTCAAAACAAGAGGATTTGTTGCCGCGAGCCCTCATC	3602	3543	AAGGGGACTCAGCTCTCAGGAGGTCAAAACAAGAGGATTTGTTGCCGCGAGCCCTCATC	3602
Db	AAGGGGACTCAGCTCTCAGGAGGTCAAAACAAGAGGATTTGTTGCCGCGAGCCCTCATC	3602	3543	AAGGGGACTCAGCTCTCAGGAGGTCAAAACAAGAGGATTTGTTGCCGCGAGCCCTCATC	3602
Qy	GlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnSerGlu	992	973	GlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnSerGlu	992
Db	GlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnSerGlu	992	973	GlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnSerGlu	992
Qy					

RESULT 15

US-09-866-866A-5
; Sequence 5, Application US/09866866A
; Patent No. US20020102244A1
; GENERAL INFORMATION:
; APPLICANT: Sorrentino, Brian
; APPLICANT: Schuetz, John
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
; FILE REFERENCE: 1340-1-021CIP2
; CURRENT APPLICATION NUMBER: US/09/866,866A
; CURRENT FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 09/584,586
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: PCT/US99/11825
; PRIOR FILING DATE: 1999-05-27


```
Qy 986 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 1005
    |||||  ::|||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 3607 GCTCTGGATACAGAAAGTGAAGAGTTGTCCAGGAAGCGCTGGACAAAGCCAGGAAGGC 3666
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy 1006 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 1025
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 3667 CGCACCTGCAATTGTGATCGCTCACCGCTGTCCACCATCCAGAACGCGGACTTGTATCGTG 3726
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy 1026 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 1045
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 3727 GTGATTGAGAACGGCAAGTCAAGGACACGGCACCCACCAGCAGCTGCTGGCGCAGAAG 3786
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy 1046 AspIleTyrPheLysLeuValAsnAla 1054
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 3787 GGCATCTACTTCTCAATGGTCCAGGCT 3813
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
```

Search completed: March 31, 2003, 16:04:42
Job time : 638.104 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 31, 2003, 13:52:37 ; Search time 444.798 Seconds
(without alignments)
2337.888 Million cell updates/sec

Title: US-09-873-409-5
Perfect score: 6159
Sequence: 1 MILGILASLVNGACPLMPL.....QELLENRDIYFKLVNAQSVQ 1222

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 574371 seqs, 425486471 residues

Total number of hits satisfying chosen parameters: 1148742

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DEV=xlh
-O=/cgn2_1/USPTO.spool/US09873409/runat_27032003_115424_19290/app.query.fasta_1.7544
-DB=PublishedApplications NA -QWTF=fastcap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=5 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09873409 @CGN 1.1 960 @runat_27032003_115424_19290
-NCPU=6 -ICPU=3 -NO_XLPHY -NO_MWAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

- 1: /cgn2_6/prodata/2/pubpna/US07_PUBCOMB.seq:
- 2: /cgn2_6/prodata/2/pubpna/PCT_NEW_PUB.seq:
- 3: /cgn2_6/prodata/2/pubpna/US06_NEW_PUB.seq:
- 4: /cgn2_6/prodata/2/pubpna/US06_PUBCOMB.seq:
- 5: /cgn2_6/prodata/2/pubpna/US07_NEW_PUB.seq:
- 6: /cgn2_6/prodata/2/pubpna/PCTU5_PUBCOMB.seq:
- 7: /cgn2_6/prodata/2/pubpna/US08_NEW_PUB.seq:
- 8: /cgn2_6/prodata/2/pubpna/US08_PUBCOMB.seq:
- 9: /cgn2_6/prodata/2/pubpna/US09_NEW_PUB.seq:
- 10: /cgn2_6/prodata/2/pubpna/US09_PUBCOMB.seq:
- 11: /cgn2_6/prodata/2/pubpna/US10_NEW_PUB.seq:
- 12: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:
- 13: /cgn2_6/prodata/2/pubpna/US60_NEW_PUB.seq:
- 14: /cgn2_6/prodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6157	100.0	3702	10	US-09-873-409-13
2	5988.5	97.2	3621	10	US-09-873-409-14
3	5328	86.5	3177	10	US-09-873-409-12
4	4561.5	74.1	2856	10	US-09-873-409-10

5	3530.5	57.3	3860	10	US-09-866-866A-1
6	3525.5	57.2	4317	9	US-10-044-671-1
7	3521.5	57.2	3860	10	US-09-866-866A-3
8	3521.5	57.2	8630	10	US-09-306-417-1
9	3521.5	57.2	8630	10	US-09-306-417-2
10	3517	57.1	4425	10	US-09-769-097-3
11	3515	57.1	4369	10	US-09-769-097-1
12	3514	57.1	4643	9	US-10-072-621-2
13	3503	56.9	4189	10	US-09-866-866A-5
14	3499	56.8	4788	10	US-09-866-866A-7
15	3481	56.5	3924	10	US-09-880-107-2299
16	3456	56.1	2066	10	US-09-873-409-9
17	3453	56.1	4254	10	US-09-917-800A-1424
18	3405	55.3	3912	10	US-09-917-800A-1560
19	2811.5	45.6	5010	10	US-09-917-800A-483
20	2697	43.8	2021	10	US-09-873-409-15
21	2528.5	41.1	1940	10	US-09-873-409-16
22	2358	38.3	4175	10	US-09-749-340-3
23	2328	37.8	3861	9	US-09-938-842A-2237
24	2259	36.7	4002	9	US-09-758-828-1
25	2259	36.7	4002	9	US-09-758-828-3
26	2129.5	34.6	3861	9	US-09-938-842A-263
27	2039.5	33.1	3512	10	US-09-749-340-7
28	1987	32.3	4653	9	US-10-101-388-2
29	1636	26.6	3792	9	US-09-882-694-10
30	1617	26.3	2698	10	US-09-749-340-5
31	1538.5	25.0	2681	10	US-09-749-340-8
32	1526.5	24.8	8036	9	US-10-101-388-1
33	1501.5	24.4	3999	9	US-09-882-694-9
34	1101.5	17.9	1175	10	US-09-873-409-11
35	1080	17.5	1810	10	US-09-749-340-4
36	1066.5	17.3	7296	10	US-09-070-927A-59
37	1027	16.7	3741	9	US-09-738-626-1251
38	1008.5	16.4	9047	10	US-09-070-927A-12
39	981.5	15.9	12438	10	US-09-070-927A-173
40	930.5	15.1	640681	10	US-09-790-988-1
41	902.5	14.7	3601	10	US-09-070-927A-255
42	901.5	14.6	6415	10	US-09-070-927A-275
43	770.5	13.5	2298	9	US-10-156-239-18
44	770.5	12.5	2298	10	US-09-795-693-18
45	770.5	12.5	3408	9	US-10-156-239-16

ALIGNMENTS

RESULT 1
US-09-873-409-13
; Sequence 13, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
; APPLICANT: Frank, Markus
; APPLICANT: Sayegh, Mohamed
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873,409
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 13
; LENGTH: 3702
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: No. US20020037522A1e
; LOCATION: (723)..(723)
; OTHER INFORMATION: n at position 723 represents any nucleotide (A, T, C or G)
US-09-873-409-13

Alignment Scores:
Pred. No.: 0
Score: 6157.00
Percent Similarity: 100.00%
Length: 3702
Matches: 1222
Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.97% Indels: 0
DB: 10 Gaps: 0

US-09-873-409-5 (1-1222) x US-09-873-409-13 (1-3702)

QY 1 MetileuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
DB 34 ATGATCTGGGTATACCTGGCATCTACTGGTCAATGGAGCGCTTCTTAAAGCCACTG 93
QY 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
DB 94 GTTTTAGCAAAATACGATGATTAACCTTATTAGTGGATGCTAGTCCAAACCTAACACATAC 153
QY 41 SerPhePheArgLeuThrLeuTyrValGlyIleGlyValAlaLeuIlePheGly 60
DB 154 TCATTCTTCAGGTGGACCTGTATATTAGTGGATAGGTGTGCTGCTTGAATTTTGGT 213
QY 61 TyrlleGlnIleSerLeuTyrIleThrAlaAlaArgGlnThrLysArgIleArgLys 80
DB 214 TACATACAGATTTCTTGTGATTAATTAACCTGCAGCACAGACAGAGAGATTCGAAA 273
QY 81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGly 100
DB 274 CAGTTTTTTTCAATTCAGTTTTGGCACAGACATCGGCTGGTTGATAGCTGTGACATCGGT 333
QY 101 GluLeuAsnThrArgMetThrAspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
DB 334 GAACCTTAACACTCGCATGACAGACATTTGAACAAATCAGTGTATGAGATAAGATT 393
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
DB 394 GCTCTGTGTTTCCAAACATGCTACTTTTTTCGATTTGGCTGGCAGTTGTTGGTGAAG 453
QY 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
DB 454 GGCTGGAAACTCACCTTAGTGACTCTATCCAGCTCTCTCTTATATAGTGTTCAGCGCA 513
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
DB 514 GCATGTTCTAGGATGGTCATCTCATTCACAGTAAGGAATTAAGTGCTTATCCAAAGCT 573
QY 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
DB 574 GGGCTGTGGCAGAGAGAGCTTGTGTCATCAATCCGACACATCATAGCTTTAGGGCCAG 633
QY 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaThrPheTyrPhe 220
DB 634 GAGAAAGAACTTCAAGGCTCTTTCTTTTAAATATATAACAAGATATGCTTGGTTTATTTT 693
QY 221 ProGlnTrpLeuLeuSerCysValLeu**PheValArgTyrThrGlnAsnLeuLysAsp 240
DB 694 CCCAGTGGCTTACTAAGTGTGTTCTGTTNTTGTAAAGGTATACAGAAATCTCAAGAT 753
QY 241 AlalysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
DB 754 GCAAGGATTTGGCATAAAAGACTATAGCTTCAAAAGTGCTCTTGGTGTGTGATC 813
QY 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
DB 814 TTTCTTTATGAATGAACCTATGGAAGTGTGCTTTTGGTATGGAACCTCTTGAATCTTAAT 873
QY 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
DB 874 GGAGAACTTGATATACCATCGGAGCTGTTCTTGTGCTGTTTCTTGTAGTAAATCCATAGC 933
QY 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
DB 934 AGTATTGCAATGGAGCAGCAGCTCCCTCACTTTGAAACCTTCCGAATAGCCCGAGAGCT 993
QY 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
DB 994 GCCTTTTCATATTTCCAGGTATTATGATAAGAAACCCAGTATAGATAACTTTTCCACAGCT 1053

QY 341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
DB 1054 GGATATAAACCTGATCCATAGAGGAAGTCTGGAAATTTAAAAATGTTCTTTCATATTAT 1113
QY 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
DB 1114 CCATCAAGACCATCTATCAAGATTCTGAAAGGTCTGAAATCTCAGAATTAAGTCTCGAGAG 1173
QY 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuGln 400
DB 1174 ACAGTCCGCTTGGTCCGCTCAATGGCAGTGGAGAGTACGGTAGTCCAGCTTCTGCAG 1233
QY 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
DB 1234 AGGTTATATGATCCGATGATGGCTTATCATGGTGATGAGATGACATCAGAGCTTTA 1293
QY 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
DB 1294 AATGTCCGCAATTCGAGACCATATTTGGAGTGGTTAGTCAAGAGCTGTTTGTTCGGG 1353
QY 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
DB 1354 ACCACCATCAGTAACAAATATCAAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAG 1413
QY 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
DB 1414 AGACGACGACGAGGACCAATGCGTATGATTTTATCATGGAGTTTCTTAATAATTTAAT 1473
QY 481 ThrLeuValGlyLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
DB 1474 ACATTGGTAGGGGAAAAGAGAGCTCAAAATGAGTGGAGGCGCAGAAACAGAGGATCGCAAT 1533
QY 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
DB 1534 GCTCGTGGCTTAGTTCGAAACCCCAAGATTTCTGATTTTAGATGAGGCTACGTCGCTCG 1593
QY 521 AspSerGluSerLysSerAlaValGlnAlaLeuGluLysAlaSerLysGlyArgThr 540
DB 1594 GATTCAAGAAAGCAAGTCACTGTTCAAGCTGCTGAGGAGGCGGAGCAAGGTCGGACT 1653
QY 541 ThrIleValAlaAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560
DB 1654 ACAATCGTGGTAGCACACGACTTCTACTATTTCGAAGTGCAGATTTGATTTGACCCCTA 1713
QY 561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlalysArgGlyLeu 580
DB 1714 AAGGATGGAAATGCTGGCGAGAAAGGAGCAGCATGCTGAACTAATGGCAAAACGAGTCTA 1773
QY 581 TyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMet 600
DB 1774 TATTATTCCTTGTGATGTCACAGGATATTTAAAAAGCTGTGAACAGATGGAGTCAATG 1833
QY 601 ThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLys 620
DB 1834 ACATATTCCTACTGAAAGAAAGACCAACTCACTCTCTCTGCACTCTGTGAAGAGATCAAG 1893
QY 621 SerAspPheIleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGlu 640
DB 1894 TCAGATTCATTGACAAAGCTGAGGAATCCCAACTATAAAGAGATAAGTCTTCTCTGAA 1953
QY 641 ValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGly 660
DB 1954 GTCTCTCTATTAATAAATTTTAAAGTTAAACAGCTGATGGCTTTTGTGTTCTGGGG 2013
QY 661 ThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLys 680
DB 2014 ACATTGGCTTCTGTTCTAAATGGAAGTGTTCATCCAGTATTTTCCATCATCTTTGCAAAA 2073
QY 681 IleIleThrMetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyr 700
DB 2074 ATTATAACCATGTTTGGAAATTAATGATAAACCACATTAAGCATGATGAGAAATTTAT 2133


```
QY 701 SerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeu 720
DB 2134 TCATGATATTCGTCATTTGGGTGTTATTTGCTTTGTCAGTTATTTATTCATGCGAGGATTA 2193
QY 721 PheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAla 740
DB 2194 TTTTACGGCAGACGAGGGGAAATTTTAAACGATGAGATTAAACACTTTGGCCCTTCAAGCC 2253
QY 741 MetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThr 760
DB 2254 ATGTTATATCAGGATATTCGCTGTTTGTATGAAAGGAAACACAGCAGAGGCTTGACA 2313
QY 761 ThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyVal 780
DB 2314 ACAATATTAGCCATAGATATAGCACAAATTTCAAGGAGCAACAGGTTCCAGGATTGGCGTC 2373
QY 781 LeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyr 800
DB 2374 TTAACACAAATGCAACTAACATGGGACTTTTCAGTTATCATTTCCCTTTATATATGATGG 2433
QY 801 GluMetThrPheLeuIleLeuSerIleAlaProValIleAlaValThrGlyMetIleGlu 820
DB 2434 GAGATCACATTCCTGATTCTGAGTATTGCTCCAGTACTTGGCGTGACAGGATGATTGAA 2493
QY 821 ThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLys 840
DB 2494 ACCGACGAATGACTGGATTTGCAACAAAGATAAGCAAGAACTTAAGCATGCTGGAAAG 2553
QY 841 IleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAla 860
DB 2554 ATAGCAACTGAAAGCTTTGGAGAAATATACGTACTATATAGTGTCTATTAACAGGGGAAAGCC 2613
QY 861 PheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAla 880
DB 2614 TTCGACCAATGATGAGAGAGATGCTTTCAGACTCAACACAGAAATACCTCGAAGAAGCA 2673
QY 881 GlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAla 900
DB 2674 CAGATTATTGGAAGCTGTTATGCTATTCAGCTATTCAGCCATGCTTTATATATTTGCTATGCAGCA 2733
QY 901 GlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPhe 920
DB 2734 GGGTTTCGATTGGAGCCCTATTTAAATTCAGCTGGACGAAATGACCCAGAGGCGCATGTTTC 2793
QY 921 IleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAla 940
DB 2794 ATAGTTTTTACTGCATTTGCATATGAGCTATGCGCATCGGAAACAGCTCGTTTGGCT 2853
QY 941 ProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLys 960
DB 2854 CCTGAATATTCCAAAGCCAAATCGGGGCTGCGCATCTGTTGCTTTGTTGGAAGAAAGAA 2913
QY 961 ProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeu 980
DB 2914 CCAAAATATAGACAGCCGAGTCAAGAAGGGAAGGAAACCCAGACACATGTGAAGGGAATTTA 2973
QY 981 GluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGly 1000
DB 2974 GAGTTTCGAGAAGTCTCTTCTCTATTCATTCATGTCGCCAGATGTTTTCATCTCCGTGGC 3033
QY 1001 LeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGly 1020
DB 3034 TTATCCCTCAGTATTGCGGAGGGAAGACAGTAGCATTTGTGGGGAGCAGCGCTGGG 3093
QY 1021 LysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeu 1040
DB 3094 AAAAGCACTTCTGTTCAACTTCTGCAGAGACTTTATGACCCCGTCGAAGGACAAAGTGCTG 3153
QY 1041 PheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIle 1060
DB 3154 TTTGATGGGTGGATGCAAAAGAAATTTGAATGTACAGTGGCTCCGTTCCCAAAATAGCAATC 3213
QY 1061 ValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAsp 1080
```

```
DB 3214 GTTCCTCAAGAGCCCTGTGCTCTTCAACTGCAGCATTGTGAGAACATCGCCCTATGGTGAC 3273
QY 1081 AsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHis 1100
DB 3274 AACAGCCGTGTGGTCCATTAGATGAGATCAAGAAGCGCAANTGCAGCAATATCCAT 3333
QY 1101 SerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGln 1120
DB 3334 TCTTTTATTGAAGGCTCTCCCTGAGAAATACAAACACACAAAGTTGGACTGAAAGGAGCACAG 3393
QY 1121 LeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLys 1140
DB 3394 CTTTCTGGCGCCAGAAACAAAGACTAGCTATTGCAAGGGCTCTTCTCCAAAACCCAAA 3453
QY 1141 IleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGln 1160
DB 3454 ATTTTATTGTTGGATGAGCCACTTCAGCCCTCGATATGACAGTGAAGAGGTGGTTCAG 3513
QY 1161 HisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSer 1180
DB 3514 CATGCCCTTGATAAAGCCAGGACGCGGAGGACATGCTAGTGGTCACTCACAGGCTCTCT 3573
QY 1181 AlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGly 1200
DB 3574 GCAATTTCAGAACCGAGATTGATAGTGGTCTTGCACAATGGAAAGATAAAGGAACAAGGA 3633
QY 1201 ThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
DB 3634 ACTCATCAAGAGCTCTGAGAAATCGAGACATATATTTTAAGTTAGTGAATGCACAGTCA 3693
QY 1221 ValGln 1222
DB 3694 GTGCAG 3699
RESULT 2
US-09-873-409-14
; Sequence 14, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
; APPLICANT: Frank, Markus
; APPLICANT: Sayegh, Mohamed
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
; FILE OF INVENTION: Homologue on Chromosome 7p15-21 and Uses Thereof
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873,409
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
; LENGTH: 3621
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-873-409-14
Alignment Scores:
Pred. No.: 0 Length: 3621
Score: 5988.50 Matches: 1195
Percent Similarity: 97.79% Conservative: 0
Best Local Similarity: 97.79% Mismatches: 0
Query Match: 97.23% Indels: 27
DB: 10 Gaps: 1
US-09-873-409-5 (1-1222) x US-09-873-409-14 (1-3621)
QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
DB 34 ATGATCCTGGGTATACCTGGCATCTGGCTCAATGGAGCTGCTCTCTTTAATGCGCACTG 93
QY 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
DB 94 GTTTTAGGAGAAATGAGTGATAACCTTATTAGTGGATGTCTAGTCCAAACTAACACATAC 153
```


41 SerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGly 60
Db TCCTTCTTCAGGTTGACCTGTATTATGTTGGATAGGTCTGCTGCCCTGATTTTGGT 213
61 TyrIleGlnIleSerLeuThrIleIleThrAlaAlaArgGlnThrIleArgLys 80
Db TACATACAGATTTCCTTGTGGATTATTAACATCGCAGCACAGACAGAGAGATTCCGAAA 273
81 GluPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGly 100
Db CAGTTTTTTTCATTGAGTTTGGCAGACAGACATCGCGTGGTTGATAGCTGTGACATCGGT 333
101 GluLeuAsnThrArgMetThrAspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db GAACCTTAACACTCGCATGACAGACATTTGACAAATCAGTGTGATTTGGATTAAGATT 393
121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db GCTCTGTGTTTCAAAACATGCTACTTTTTCGATTGGCCCTGGCAGTTGTTGGTGAAG 453
141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db GGCTGGAAACTCACCTTAGTACTCTATCCACGCTCTCTCTTAAATGGCTTCAGCGGCA 513
161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db GCATGTTCTAGGATGGTCACTCATTTGACCAGTAAGNATTAAGTGCCTATTTCCAAAGCT 573
181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db GGGGCTGTGGCAGAGAAGCTTGTGTCATCAATCCGAACAGTCATAGCCTTTAGGGCCAG 633
201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
Db GAGAAAGAACTTCAA----- 648
221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
Db -----AGGTATACACNGAATCTCAAGAT 672
241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
Db GCAAGAGATTTTGGCATAAAGAGACTATAGCTTCAAAAGTGCTCTTGGTGTGTGATC 732
261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
Db TTCTTTATGAATGGAACTATGAGCTTGCCTTTTGGTATGGAACCTCTTGATTTCTTAAT 792
281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
Db GGAGAACTCGATATACCATCGGACCTGTTCTTGTGTTTCTTTAGTGTAAATCCATAGC 852
301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
Db AGTTATTGCAATTGGAGCAGCAGCTCCCTCACTTTGAAACCTTCGCAATAGCCCGAGGAGCT 912
321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
Db GCCTTTTCATATTTTCCAGGTTATTGATAAGAAACCCCATAGATATAGATTTTCCACAGCT 972
341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
Db GGATATAAACCCTGATNCCATAGAGAACTGTGGAAATTTAAAGATGTTCTTCAATTAT 1032
361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
Db CCATCAAGACCATCTATCAAGATTTCTCAAGAGTCTGAATCTCAGAAATTAAGTCTGGAGAG 1092
381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
Db ACAGTCCCTTGGTGGCTCTCAATGGCAGTGGGAAGATAGCGTAGTCCAGCTTCTGCAG 1152
401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420

1153 AGGTTATATGATCCGATGATGGCTTTTATCATGGTGGATGAGAAATGACATCAGAGCTTTA 1212
421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
Db AATGTGCGGCATTTATCGAGACCATATTTGGAGTGGTTAGTCAAGAGGCTGTTTGTTCGGG 1272
441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspValThrAspGluGluMetGlu 460
Db ACCACCATCAGTAACATATCAAGTATGCGCAGATGATGTGACTGATCGAGAGATGGAG 1332
461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
Db AGACGACGAGGGAAGCAAAATGCTATGATTTTATCATGGAGTTTCTTAATAAATTAAT 1392
481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
Db ACATTTGGTAGGGGAAAAGAGGCTCAAAATGAGTGGAGGCGCAAAACAGAGGATTCGCAATT 1452
501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
Db GCTGTGCTTGTAGTTGCAAAACCCCAAGATTTGATTTTAGATGAGGCTACGTTGCGCTG 1512
521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThr 540
Db GATTACAGAAAGCAAGTCAAGCTGTTCAAGCTGCACCTGGAGAGCGCAAGAGTGGACT 1572
541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560
Db ACAATCGTGTAGCACACCGACTTTCTACTATTTCGAAGTGCAGATTTGATTTGACCCCTA 1632
561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580
Db AGGATGGAATGCTGGCGGAGAAAGGAGCAGCATGCTGAACCTAATGGCAAAACGAGGTCTA 1692
581 TyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMet 600
Db TATTATTCACTTGTGATGTCACAGGATATTTAAAAAGCTGATGAACAGATGCAATG 1752
601 ThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLys 620
Db ACATATTTCTACTGAAAGAAAGACCAACTCACTCTCTCTGCTCTGCTGTGAAGACATCAAG 1812
621 SerAspPheIleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGlu 640
Db TCAGACTTCATTGACAGGCTGAGGATCCACCAATCTTAAGAGATAGTCTTCTCTGAA 1872
641 ValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpPheValValLeuGly 660
Db GTCTCTCTATTTAAATAATTTTAAAGCTTAAACAGCTGAAATGGCTTTTGTGCTTCTGGGG 1932
661 ThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLys 680
Db ACATTTGGCTTCTGTTCTAAATGGAACTGTTCATCCAGTATTTTCCATCATCTTTGCAAAA 1992
681 IleIleThrMetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyr 700
Db ATTATTAACCATGTTTGGAAATTAATGATAAAACACCATTAAGCATGATGACAGAAATTTAT 2052
701 SerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeu 720
Db TCCATGATATTCGTCAATTTGGGTGTTATTGCTTTGTGCTTATTTTCATGCGAGGATTA 2112
721 PheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAla 740
Db TTTTACGCGCAGACGAGGGGAAATTTTAAACGATGAGATTAAACACTTGGCCTTTCAAGGCC 2172
741 MetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThr 760
Db ATGTTATATCAGGATATTCCTGCTGTTTGTATGAAGAAAGAAACAGACAGAGGCTTGACA 2232
761 ThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyVal 780


```
Db 2233 ACAATATTAGCCATAGATATAGCACAAATTCAGGAGCAACAGGTTCCAGGATTGGCGTC 2292
Qy 781 LeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyr 800
Db 2293 TTAACACAAATGCAACTAACATGGGACCTTTCAAGTATCATCTTCTTATATATGATGG 2352
Qy 801 GluMetThrPheLeuLeuLeuSerIleAlaProValLeuAlaValThrGlyMetIleGlu 820
Db 2353 GAGATGACATTCCTGATCTCTGAGTATTGCTCCAGTACTTGCCTGACAGGAATGATTGAA 2412
Qy 821 ThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLys 840
Db 2413 ACCGCAATGATGATGATTTGCAACAAAGATAAGCAAGAACTTTAAGCATGCTGGAAG 2472
Qy 841 IleAlaThrGluAlaLeuGluAsnIleAtrThrIleValSerLeuThrArgGluLysAla 860
Db 2473 ATAGCAACTGAAGCTTTGGAGATATACGTACTATATGTTGCTATTAAACAGGGAAGGCC 2532
Qy 861 PheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAla 880
Db 2533 TTCGAGCAAAATGATGAAGAGATGCTTTCAGACTCAACACAGAAATACCTCGAAGAAAGCA 2592
Qy 881 GlnIleLeuGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAla 900
Db 2593 CAGATTATTGGAAGCTGTTATGATTCAGCCATGCTCTTATATATTGCTCATGACGCA 2652
Qy 901 GlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPhe 920
Db 2653 GGGTTTCGATTTCGAGCTATTATTAATCAAGCTGGAGATGATCCACAGAGGCGATGTC 2712
Qy 921 IleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAla 940
Db 2713 ATAGTTTCTACTGCAATTCATATGAGCTATGCGCATGCGCAATCGTCTGTTTGGCT 2772
Qy 941 ProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuLysLys 960
Db 2773 CCTGAATATTCGAAGCCAAATCGGGGGCTGCGCATCTGTTGCTTTGGAAAGAAA 2832
Qy 961 ProAsnIleAspSerArgSerGlnGluGlyLysProAspThrCysGluGlyAsnLeu 980
Db 2833 CCAATATAGACAGCGCATCAAGAGGGGAAAGCCACACATGTGAAGGGAATTTA 2892
Qy 981 GluPheArgGluValSerPheTyrProCysArgProAspValPheIleLeuArgGly 1000
Db 2893 GAGTTTCGAGAAGTCTCTTCTCTATCCATGTCGCCAGATGTTTTCATCTCCGTGGC 2952
Qy 1001 LeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGly 1020
Db 2953 TTTATCCCTCAGTATTGAGCGAGGAAGACAGTAGCATTTGTGGGAGACGCGCTGTGGG 3012
Qy 1021 LysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeu 1040
Db 3013 AAAAGCACTTCTGTTCAACTTCTGCAGAGACTTTATGACCCCGTGCAGAGCAAGTGCTG 3072
Qy 1041 PheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIle 1060
Db 3073 TTTGATGGTGTGATGCAAAAGAAATGAATGTACAGTGGCTCCGTTCCCAAAATAGCAATC 3132
Qy 1061 ValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAsp 1080
Db 3133 GTTCCCTCAAGAGCGCTGTGCTCTTCACTGAGCATTCGTCAGAACATCGCTATGTTGGTAC 3192
Qy 1081 AsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHis 1100
Db 3193 AACAGCGTGTGGTGCCATTAGATGAGATCAAGAGCCGCAATGACGCAATATATCCAT 3252
Qy 1101 SerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGln 1120
Db 3253 TCTTTTATTGAAGGTCTCCCTGAGAAATACACACACAAGTTGACTGAAAGGAGCACAG 3312
Qy 1121 LeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuGlnLysProLys 1140
Db 3313 CTTTCTGGCGGCCAGAAACAAGACTAGCTATTGCAAGGGCTCTTCTCCAAAACCCAAA 3372
```

```
Qy 1141 IleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGln 1160
Db 3373 ATTTTATTGTTGGATGAGGCCACTTCAGCCCTCGATAATGACAGTGAAGAGTGGTTTCAG 3432
Qy 1161 HisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSer 1180
Db 3433 CATGCCCTTTGATAAAGCCAGGACGGGAAGACATGCTTAGTGGTCACTCACAGGCTCTCT 3492
Qy 1181 AlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGly 1200
Db 3493 GCATTCAGAACGCGAGATTTGATAGTGGTTCTGCACAAATGGAAGATAAAGGAACAAGGA 3552
Qy 1201 ThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
Db 3553 ACTCATCAAGAGCTCTCGAGAAATCGAGACATATATTTTAAAGTTAGTGAATGCACAGTCA 3612
Qy 1221 ValGln 1222
Db 3613 GTGCAG 3618
RESULT 3
US-09-873-409-12
; Sequence 12, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
; APPLICANT: Frank, Markus
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
; TITLE OF INVENTION: Homologue on Chromosome 7p15-21 and Uses Thereof
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873,409
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 3177
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: No. US20020037522A1e
; LOCATION: (198)..(198)
; OTHER INFORMATION: n at position 198 represents any nucleotide (A, T, C or G)
US-09-873-409-12
Alignment Scores:
Pred. No.: 0 Length: 3177
Score: 5328.00 Matches: 1058
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 86.51% Indels: 0
DB: 10 Gaps: 0
US-09-873-409-5 (1-1222) x US-09-873-409-12 (1-3177)
Qy 165 MetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAla 184
Db 1 ATGGTCACTCTCATTCACCAAGTAAGGAATTAAGTGCCTATTCCAAAGCTGGGGCTGTGGCA 60
Qy 185 GluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeu 204
Db 61 GAAGAAGTCTTGTCATCAATCCGAACAGTCATAGCCCTTTAGGGCCCGAGGAAGAAGACTT 120
Qy 205 GlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeu 224
Db 121 CAAAGGCTCTTCCTTTTAAATATAACAGATATGCTTGTTTATTTTCCCCAGTGGCTA 180
Qy 225 LeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPhe 244
Db 181 CTAAAGTTGTTCTGTTTNTTGTAAAGTATACACAGAACTCAAGAGATGCAAGAGGATTTT 240
Qy 245 GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsn 264
```


Db 241 GGCATAAAAAAGGACATAGCTCAAAAGTCTCTCTGGTGTGTACTCTTTTATCAAT 300
Qy 265 GlyThrTyrGlyLeuAlaPheTyrGlyThrSerLeuIleLeuAsnGlyGluProGly 284
Db 301 GGAACTTATGGACTTGTCTTTTGGTATGGAACCTCTCTTGATTTCTTAATGAGAACCTTGA 360
Qy 285 TyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIle 304
Db 361 TATACCAATCGGACTGTCTCTGCTGTTTTCTTAGTGTATATCCATAGCAGTATTGCATT 420
Qy 305 GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 324
Db 421 GGAGCAGCAGTCCCTCACTTTGAAACCTTCGCAATAGCCGAGAGCTGCTTTCATATT 480
Qy 325 PheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysPro 344
Db 481 TTCAGGTTATTGTATAAGAACCCAGATATAGATAAATTTTCCACAGCTGATATAACCT 540
Qy 345 GluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgPro 364
Db 541 GAATCCATAGAGGAACGTGTGGAATTTAAAAATGTTCTTTCAATTATCCATCAAGACCA 600
Qy 365 SerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu 384
Db 601 TCTATCAAGATTCTGAAAGGTCTGAATCTCAGAAATTAAGTCTGGAGAGACAGTCGCCCTG 660
Qy 385 ValGlyLeuAsnGlySerGlyLysSerThrValIleGlnLeuLeuGlnArgLeuTyrAsp 404
Db 661 GTCGGTCTCAATGGCAGTGGGAAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGAT 720
Qy 405 ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis 424
Db 721 CCGGATGATGGCTTTATCATGGTGGATGAGATGACATCAGAGCTTTAAATGTGGGCAT 780
Qy 425 TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrIleSer 444
Db 781 TATCGAGACCATATTGGAGTGGTGTAGTCAAGAGCGCTGTTTGTTCGGGACCAACCATCAGT 840
Qy 445 AsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArg 464
Db 841 AACAAATATCAAGTATGACGAGATGATGTGACTGATGAAGAGATGGAGAGCAGCAAGG 900
Qy 465 GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly 484
Db 901 GAACCAATCGGTATGATTTATCATGGAGTTCTCTAATAAATTAATACATTTGGTAGG 960
Qy 485 GluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu 504
Db 961 GAAAAAGGAGCTCAAAATCAGTGGAGGGCAGAAAACAGAGGATCGCAATTTGCTGCGCTTA 1020
Qy 505 ValArgAsnProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspSerGluSer 524
Db 1021 GTTCGAACCCCAAGATCTGATTTTAGATGAGGCTACGCTGCCCTGGATTCAGAAAGC 1080
Qy 525 LysSerAlaValGlnAlaLeuGluLysAlaSerLysGlyArgThrThrIleValVal 544
Db 1081 AAGTCAGCTGTTCAAGTGCCTCGAGAGCGCAGCAAGGTCGAGTCAATCGTGTA 1140
Qy 545 AlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMet 564
Db 1141 GCACACCGACTTTCTACTATTTCGAAGTGCAGATTGTGATGACCCCTAAAAGGATGGAATG 1200
Qy 565 LeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeu 584
Db 1201 CTGGCGGAGAAAGGAGCAGACATGCTGGAATATGTCGCAAAACAGGCTCATATTATTCATT 1260
Qy 585 ValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThr 604
Db 1261 GTGATGTTCACAGGATATTAATAAGCTGATGACAGATGGNGTCAATGACATATTCTACT 1320
Qy 605 GluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle 624
Db 1321 GAAAGAAAGACCAACTCACTTCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTCATT 1380

Qy 625 AspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeu 644
Db 1381 GACAAGCTGAGGAATCCACCCAACTAAAGAGATAAGTCTTCTCTGAAGTCTCTCTATTATTA 1440
Qy 645 LysIleLeuLysLeuAsnLysProGluTyrProPheValValLeuGlyThrLeuAlaSer 664
Db 1441 AAAATTTTAAAGTTAAACAAGCTGAATGCCCTTTGTGGTCTCTGGGACATTTGGCTTCT 1500
Qy 665 ValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMet 684
Db 1501 GTTCTAAATGGAACTGTTTCATCAGTATTTCATCATCTTTGCAAAAATTAACCAATG 1560
Qy 685 PheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePhe 704
Db 1561 TTTGGAAATAATGATANAACACACATTAAGCATGATGCAGAAATTTATTCCATGATATTC 1620
Qy 705 ValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArg 724
Db 1621 GTCAATTTTGGTGTATTGCTTTGTGCTATTATTTTCATGAGGATTAATTTACGGCAGA 1680
Qy 725 AlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGln 744
Db 1681 GCAGGGGAAATTTTAACGATGAGATTAAAGACACTTGGCCCTTCAAGGCCATGTTATATCAG 1740
Qy 745 AspIleAlaTyrPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAla 764
Db 1741 GATATTGCTGCTGTTGATGAAAAGGAAAACACACAGCAGGAGCTTGACAACTATATTAGCC 1800
Qy 765 IleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsn 784
Db 1801 ATAGATATAGCAAAATTCAGGAGCAACAGGTTCCAGGATTTGGCGCTTTAAACACAAAAT 1860
Qy 785 AlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyrGluMetThrPhe 804
Db 1861 GCAACTTAACATGGAGCTTTTCAGTTATCTTCTTATATATGATGGAGTGGAGATGACATTC 1920
Qy 805 LeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMet 824
Db 1921 CTGATTTCTGAGTATTGCTCCAGTACTTTCGCGTGACAGGAATGATGAAACCCGAGCAATG 1980
Qy 825 ThrGlyPheAlaAsnLysAspLysGlnLeuLysHisAlaGlyLysIleAlaThrGlu 844
Db 1981 ACTGGATTTGCCAACAAAGATAAGCAAGAACTTAAGCATGCTGGAAGAGATAGCAACTGAA 2040
Qy 845 AlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMet 864
Db 2041 GCTTTGGAGAAATATAGTACTATAGTGTCTATTAAACAAGGAAAAAGCCCTTCGAGCAATG 2100
Qy 865 TyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGly 884
Db 2101 TATGAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCACAGATATTGGA 2160
Qy 885 SerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPhe 904
Db 2161 AGCTGTTATGCATTCAGCCATGCTTTATATATATTTTGCCTATGTCAGCAGGGTTTCGATTT 2220
Qy 905 GlyAlaTyrIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThr 924
Db 2221 GGAGCTTATTAAATCAAGCTGGACGAATGACCCAGAGGGCATGTTTCATAGTTTACT 2280
Qy 925 AlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSer 944
Db 2281 GCATTTGCATATGGAGCTATGGCCATCGGAAAAACCGCTCGTTTGGCTCCTGAATATTC 2340
Qy 945 LysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAsp 964
Db 2341 AAAGCCAAATCGGGGCTCGCATCTGTTTGCCTTGTGGAAAAAGAAACCAATATAGAC 2400
Qy 965 SerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGlu 984
Db 2401 AGCCCGAGTCAAGAGGAAAAAGCCAGACACATGTGAAGGGGAATTTAGAGTTTTCAGAA 2460


```
QY 985 ValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSer 1004
Db 2461 GTCTCTTCTTCTATCCATGTCGCCAGAGTGTTCATCTCCGTGGCTTATCCCTCAGT 2520
QY 1005 IleGluArgGlyLeuThrValAlaPheValGlySerSerGlyCysGlyLeuSerThrSer 1024
Db 2521 ATTGACGCGAGGAAGACAGTAGCATTTGTGGGGAGCAGCGCTGTGGGAAAGCATCTTCT 2580
QY 1025 ValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyVal 1044
Db 2581 GTTCAACTTCTGCAGAGACTTTATGACCCCGTCAAGGACAAGTGTGTTGATGGTGTG 2640
QY 1045 AspAlaLysGlnLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGlu 1064
Db 2641 GATGCAAAAAGAAATTGAATGTACAGTGGCTCGCTTCCCAATAGCAATCGTTCCTCAAGAG 2700
QY 1065 ProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgVal 1084
Db 2701 CCGTGTCTCTTCAACTGACGACTTGTGAGAACATCGCCTATGGTGACAAACGCCGTGTG 2760
QY 1085 ValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGlu 1104
Db 2761 GTGCCATTAGATGAGATCAAGAGCGCGCAATGCAGCAATATCCATCTTTTATTGAA 2820
QY 1105 GlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGly 1124
Db 2821 GGTCTCCCTGAGAAATACAAACACACAAGTTGGACTGAAAGGAGCACAGCTTTCTGGCGGC 2880
QY 1125 GlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeu 1144
Db 2881 CAGAAACAAAGACTAGCTATTGCAAGGGCTCTTCTCCAAAACCCAAAATTTTATTGTTG 2940
QY 1145 AspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAsp 1164
Db 2941 GATGAGCCCACTTCAGCCCTCGATATATGACAGTGAAGAGTGTTCAGCATGCCCTTGAT 3000
QY 1165 LysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsn 1184
Db 3001 AAAGCCAGGACGGGAAGGACATCGCTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAAC 3060
QY 1185 AlaAspLeuIleValValLeuHisAsnGlyLysIleLysGlnGlnGlyThrHisGlnGlu 1204
Db 3061 GCAGATTGATAGTGGTCTTGCAATATGGAAGATGAAGGAACCAAGAACTCATCAAGAG 3120
QY 1205 LeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 1222
Db 3121 CTCTCGAAGAAATCGAGACATATATTTTAAAGTTAGTGAATGCACAGTCAGTGCAG 3174
```

RESULT 4

```
US-09-873-409-10
; Sequence 10, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
; APPLICANT: Frank, Markus
; APPLICANT: Sayegh, Mohamed
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
; TITLE OF INVENTION: Homologue on Chromosome 7p15-21 and Uses Thereof
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873,409
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
; LENGTH: 2856
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-873-409-10

Alignment Scores:
Pred. No.: 0 Length: 2856
Score: 4561.50 Matches: 912
Percent Similarity: 97.96% Conservative: 0
Best Local Similarity: 97.96% Mismatches: 1
```

```
Query Match: 74.06% Indels: 18
DB: 10 Gaps: 1
US-09-873-409-5 (1-1222) x US-09-873-409-10 (1-2856)
QY 293 ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHis-PheG1 312
Db 114 GTTTTCTTTTAGTGTATCCATAGCAGTATTTCATTTGGAGCAGCAGTCCCTCATTTATT-- 171
QY 312 uThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPr 332
Db 172 -----GATAAGAAACC 182
QY 332 oSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValG1 352
Db 183 CAGTATAGATAACTTTCCACAGCTGGATATAAACCTGAATCCATAGAAGAACTGTGA 242
QY 352 uPheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLe 372
Db 243 ATTTTAAAAATGTTTCTTTCAATTATCCATCAAGACCATCTATCAAGATTCTGAAAGGTCT 302
QY 372 uAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLy 392
Db 303 GAATCTCAGAAATTAGTCTGGAGAGACAGTCGCTTGGTTCGCTCTCAATGGCAGTGGGA 362
QY 392 sSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVa 412
Db 363 GAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCGATGATGGCTTTTATCATGCT 422
QY 412 lAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVa 432
Db 423 GGATGAGAAATGACATCAGAGCTTTTAAATGTGGCGCATTTATCGAGACCATATTTGGAGTGT 482
QY 432 lSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAs 452
Db 483 TAGTCAAGAGCCTGTTTTGTTCCGGACCACTCATGATTAACAATATCAGTATGGACGAGA 542
QY 452 pAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaIleAspPheI 472
Db 543 TGATGTGACTGATGAAGACATGGAGAGACAGCAAGGGGAAGCAAAATGCGTATGATTTTAT 602
QY 472 eMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGl 492
Db 603 CATGGAGTTTCTTAATAATTTAATACATTTGGTGGGGAAGAGGAGCTCAAAATGAGTGG 662
QY 492 vGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuI 512
Db 663 AGGCGAAGAACACAGAGATCGCAATTGCTCGTCCCTTAGTTTGGAAACCCCAAGATTTCTGAT 722
QY 512 eLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLe 532
Db 723 TTTAGATGAGGCTACGTCGCTCGCTGGATTTCAGAAAGCAAGTCAGCTGTTTCAAGTCAC 782
QY 532 uGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIleAr 552
Db 783 GGAGAAAGCGCAGCAAAAGTTCGAGCTACAATCGTGGTAGCACACCCGACTTTTCTACTATT 842
QY 552 gSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAl 572
Db 843 AAGTGCAGATTGATTGTGACCCCTAAAGGATGAATGCTGGCGGGAAGAGGACACATGC 902
QY 572 aGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIleLysLy 592
Db 903 TGNACTAATGGCAAAACGAGGTCTATATTATTCACCTTGTGATGTACAGATATTAAAA 962
QY 592 sAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuPr 612
Db 963 AGCTGATGAACAGATGAGTCAATGACATATTCTACTGAAAGAAAGACCAACTCACTTCC 1022
QY 612 oLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThrGl 632
Db 1023 TCTGCACCTCTGTGAAGAGCATCAAGTCAGACTTCATTGTGAAGGCTGAGGAATCCACCCA 1082
```


632 nSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysPr 652
1083 ATCTAAAGAGATAAGTCTCTCTGAAGTCTCTCTATTAAAAATTTAAAGTTAAACAGCC 1142
652 oGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPr 672
1143 TGAATGGCCCTTTGTGGTCTCTGGGACATTTGGCTCTCTTCTAAATGGAACTGTTCATCC 1202
672 oValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrTh 692
1203 AGTATTTTCCATCATCTTTGCCAAAATTAACCATGTTTGGAAAATTAATGATAAACCCAC 1262
692 rLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPh 712
1263 ATTAAGCATGATGCAGAAATTTATTCATGATATTCGTCAATTTGGGTGTATTGGCTT 1322
712 eValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetAr 732
1323 TGTCAAGTATTTCATGCAGGATTTATTTACGGCAGACGAGGGGAAATTTTAAACGATGAG 1382
732 gLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLyl 752
1383 ATTAAGACACTTGGCCCTTCAAGCCATGTTATATCAGATATTGCCCTGGTTTGAAGAAA 1442
752 sGluAsnSerThrGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnI 772
1443 GGAACACAGCACAGAGGCTTGACAAATATATTAGCCATAGATATAGCAAAATTCAGG 1502
772 yAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerVa 792
1503 AGCAACAGGTTCCAGGATGGCGCTTTAAACACAAAATGCAACTAACATGGGACTTTCCAGT 1562
792 lIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProVa 812
1563 TATCATTTCCCTTATATATGATGGAGATGACACTTCCGATTCGTAGTATTGCTCCAGT 1622
812 lLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLy 832
1623 ACTTCCCGTCACAGGAATGATTGAAACCGCAGCAATGACTGGATTTGCCAACAAAGATAA 1682
832 sGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIl 852
1683 GCAAGAACTTTAAGCATCTCGAAGATAGCAACTGAAAGCTTTGGAGAATATACGTACTAT 1742
852 eValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGl 872
1743 AGTGCTATTAAACAGGGAAGAAAGCCTTCGAGCAATGATGTAAGAGATGCTTCAGACTCA 1802
872 nHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAl 892
1803 ACACAGAAATACCTCGAAGAAAGCACAGATTTATTGGAAGCTGTTATGCAATTCAGCCATGC 1862
892 aPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGl 912
1863 CTTTATATATTTTCCCTATGCAGCAGGCTTCCGATTTGGAGCCCTATTTAATTAAGCTGG 1922
912 yArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAl 932
1923 ACGAATGACCCAGAGGCGATGTTCTAGTATTTTACGTGCAATTCATATGAGCTATGGC 1982
932 aIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHi 952
1983 CATCGGAAAACGCTCTTTTGGCTCTCTGATATTTCCAAAGCCAAATCGGGGCTGGCA 2042
952 sLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluLysLysLy 972
2043 TCTGTTTCCCTTGTGGAAAAGAAACCAATATATACAGCCGCGAGTCAAGAAAGGGAATAA 2102
972 sProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPheTyrProCysAr 992
2103 GCCAGACATGTGAAGGAAATTTAGAGTTTTCGAAGATGCTCTTCTTCATCCATGTCG 2162
992 gProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAl 1012

2163 CCAGATGTTTTTCATCTCCGTGGCTTATCCCTCAGTATTGAGCGGAAAGACAGTAGC 2222
1012 aPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTy 1032
2223 ATTGTGGGGAGCAGCGGCTGTGGGAAAAGACATCTCTGTTCACCTTCTCAGAGACTTTA 2282
1032 rAspProValGlnGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGl 1052
2283 TGACCCCGTGCAAGGACAGTGTCTGTTCATGGTGGATGCAAAAGAAATTTGAATGTACA 2342
1052 nTrpLeuArgSerSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIl 1072
2343 GTGGCTCCGTTCCTCAATAGCAATCGTCTCAAGAGCCCTGTCTCTTCAACTGCGACAT 2402
1072 eAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGl 1092
2403 TGTGTGAAGACATCGCTATGGTGACAAACCGCTGTGGTGCCATTTAGATGAGATCAAGA 2462
1092 uAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnTh 1112
2463 AGCCGCAATTCAGCAAAATATCCATTTCTTTTATTGAAAGTCTCCCTGAGAAATACAACAC 2522
1112 rGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAl 1132
2523 ACAAGTTGGACTGAAAGGAGCACAGCTTCTGGCGCCAGAAACAAAGACTAGCTATTGC 2582
1132 aArgAlaLeuLeuGlnLysProLysIleLeuLeuAspGluAlaThrSerAlaLeuAs 1152
2583 AGGGCTCTTCTCCAAAACCCCAAAATTTTATTGTTGGATGAGGCCACTTCAGCCCTCGA 2642
1152 pAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCy 1172
2643 TAATGACAGTGAGAGGTGGTTCAGCATGCCCTTGATTAAGCCAGGACGGGAGGACATG 2702
1172 sLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHi 1192
2703 CCTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAACCGCAGATTTTGATAGTGTCTTGCA 2762
1192 sAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTy 1212
2763 CAATGAAAGATAAAGAAACAAAGAACTCATCAAGAGCTCTCGAGAAATCGAGACATATA 2822
1212 rPheLysLeuValAsnAlaGlnSerValGln 1222
2823 TTTTAAGTTAGTGAATGCACAGTCAGTCAGCAG 2853

RESULT 5

US-09-866-866A-1
; Sequence 1, Application US/09866866A
; Patent No. US20020102244A1
; GENERAL INFORMATION:
; APPLICANT: Sorrentino, Brian
; APPLICANT: Schuetz, John
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
; FILE REFERENCE: 1340-1-021CIP2
; CURRENT APPLICATION NUMBER: US/09/866,866A
; CURRENT FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 09/584,586
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: PCT/US99/11825
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: 60/086,988
; PRIOR FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 3860
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-866-866A-1

Alignment Scores:

Pred. No.: 4,44e-314 Length: 3860
Score: 3530.50 Matches: 682
Percent Similarity: 74.00% Conservative: 246
Best Local Similarity: 54.39% Mismatches: 263
Query Match: 57.32% Indels: 63
DB: 10 Gaps: 10

US-09-873-409-5 (1-1222) x US-09-866-866A-1 (1-3860)

Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 151 ATGGTGGTGGGAACCTTGGCTGCATCATCCATCGGCTGGACCTTCTCTCATGATGCTG 210
Qy 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
Db 211 GTGTTTGGAGAAATGACAGATATCTTTGCAATGTCAGGAAATTTAGAAAGATCTGATGTCA 270
Qy 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
Db 271 AACATCACTAATAGAAAGTATATCAATGATACAGAGGTTCTTCAATGTAATCTGGAGGAAGAC 330
Qy 44 -----ArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db 331 ATGACACAGATATGCCATTATTAACAGTGAATTTGGTCTGGGGTGTGGTGTGCTGTAC 390
Qy 62 IleGlnIleSerLeuTrpIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db 391 ATTACAGTTTCAATTTGGTGGCTGGCAGCTGGAGAACAAATACACAAATTTAGAAACAG 450
Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db 451 TTTTTCATGCTATAATCGCAGGAGATAGGCTGGTTGATGTCACAGATGTGGGGAG 510
Qy 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db 511 CTTAACACCCGCTTACAGATGATGCTCTAAGATTAATGAAGTATTTGGTGCAAAAT 570
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db 571 GGAATGCTCTTCAGTCAGTGGCAACATTTTTCACCTGGGTTTATAGTAGGATTTACACGT 630
Qy 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db 631 GGTGTGAAGCTAACCTTGTGATTTTGGCCATCAGTCTCTTCTTGGACTGTGTCAGCTGT 690
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db 691 GTCGGGCAAGATATCTTCTTACTATAAAGAACTCTTAGCGTATGCAAAAGCT 750
Qy 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db 751 GGAGCAGTAGCTGAAGAGCTTTGGCAGCAATTTAGAACTGTGATTCATTTGGAGGACAA 810
Qy 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
Db 811 AAGAAAGAACTTGAA----- 825
Qy 221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
Db 826 -----AGGTACAAACAAAATTTAGAAAGAA 849
Qy 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
Db 850 GCTAAAGAAATTTGGGATAAAGAAAGCTATTACAGCCAAATTTCTATAGTGTGCTCTTC 909
Qy 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
Db 910 CTGCTCATCTATGCATCTTATGCTCTGGCCTTCTGGTATGGGACCACCTTGGTCTCTCA 969
Qy 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
Db 970 GGGGAA-----TATTCATTTGGACAAAGTACTCACTGTATTTCTTTCTGTATTAATTTGGG 1023

Qy 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
Db 1024 GCTTTTAGTGTGGACAGCATCTCCAAGCATTTGAAGCATTTGCAAAATGCAAGAGAGCA 1083
Qy 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
Db 1084 GCTTTAGAAATCTTCAAGATAATTGATAAAGCCAGTATTGACAGCTATTTCGAAGAGT 1143
Qy 341 GlyTyrLysProGluSerIleGluThrValGluPheLysAsnValSerPheAsnTyr 360
Db 1144 GGGCACAACACAGATAATATTAAAGGAAATTTGGAATTCAGAAATGTTTCACTTTCAGTTAC 1203
Qy 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
Db 1204 CCATCTCGAAAAAGAAATTAAGATCTTGAAGGCGCTGAACCTGAAGGTGAGAGTGGGAC 1263
Qy 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
Db 1264 ACGGTGGCCCTGGTTGGAAACAGTGGCTGTGGNAGAGCACACAGTCCAGCTGTATGCGAG 1323
Qy 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
Db 1324 AGGCTCTATGACCCACAGAGGGGATGGTCAGTGTGTATGGACAGGATATTAGGACCATA 1383
Qy 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
Db 1384 AATGTAAGGTTTCTACGGGAAATCATTTGGTGTGGTCAAGTCAGGAACTGTATTTGTTGCC 1443
Qy 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
Db 1444 ACCAGATAGCTGAAACCAATTCGTATGGCCGTGAAATGTCCACCATGGATGAGATTGAG 1503
Qy 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
Db 1504 AAAGCTGTCAAGAGGCCAATGCCATGACTTATCATGAACCTGCCTCATAAATTGAC 1563
Qy 481 ThrLeuValGlyLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
Db 1564 ACCCTGGTTGGAGAGAGGGGCCAGTTGAGTGTGGTGGCAGAGCAGAGGATTCGCCATT 1623
Qy 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
Db 1624 GCACGTGCCCTGGTTCGCAACCCCAAGATCCCTCGTGGATGGAGGCCACGTCAGCTTG 1683
Qy 521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThr 540
Db 1684 GACACAGAAGCGAAGCAGCTGTTGAGTGGTCTGGATGAAGCCAGAAAGGTCCGACC 1743
Qy 541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560
Db 1744 ACCATTGTGATAGTTCATCGTTTGTCTACAGTTCGTAATGCTGACGTCATCGCTGTTTC 1803
Qy 561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580
Db 1804 GATGATGGAGTCAATTGTGGAGAAAGGAAATCATGATGAACCTCATGAAGAGAAAGCAAT 1863
Qy 581 TyrTyrSerLeuValMetSerGln-----AspIleLysLysAlaAsp 594
Db 1864 TACTTCAAACTTGTCAATGCAGACAGCAGGAAATGAAGTTGAATTAGAAATGCGACT 1923
Qy 595 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis 614
Db 1924 GATGAATCCAAAAGTGAATTCATGCTTTGGAAATGCTCTTCAATGATTCAAGATCCAGT 1983
Qy 615 SerVal-----LysSerIleLys---SerAspPheIleAspLysAla 627
Db 1984 CTAATAAGAAAAAGATCAACTCGTAGGAGTGTCCGTGGATTCACAGCCCAAGACAGAAAG 2043
Qy 628 GluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeu 647
Db 2044 CTTAGTACCAAGAGGCTCTGGATGAAGATATACCTCCAGTTTCCTTTGGAGGATATAG 2103
Qy 648 LysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsn 667


```
Db 2104 AAGCTAAATTAAGTAATGCCTTATTTGTTGGTGTATTTTGTGCATTAATAAT 2163
Qy 668 GlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn 687
Db 2164 GGAGGCTGCACACGACGATTTGCATAATATTTTCAAGAAATATAGGGGTTTTACAGA 2223
Qy 688 ---AsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIle 706
Db 2224 ATTGATGATCCTGAACCAAAACGACAGAAATAGTAACCTTGTTCACATATTTGTTCAGCC 2283
Qy 707 LeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGly 726
Db 2284 CTTGGAATATTTCTTTTATTACATTTTCTTCAAGGTTTTCATTTGGCAAAAGCTGA 2343
Qy 727 GluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIle 746
Db 2344 GAGATCCTCACCAAGCGGCTCCGATACATGATGTTTCCGATCCATGCTCAGACAGGATGTG 2403
Qy 747 AlaTyrPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAsp 766
Db 2404 AGTTGGTTTGATGACCTTAAACACACACTGGAGCATTTGACTACCGGCTCGCCAAATGAT 2463
Qy 767 IleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThr 786
Db 2464 GCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTAATTAACCCAGAATATAGCA 2523
Qy 787 AsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyrGluMetThrPheLeuIle 806
Db 2524 AATCTGGGACAGGAATAATATATATCTTCATCTATGTTGGCACTTAACACTGTTTACTC 2583
Qy 807 LeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGly 826
Db 2584 TTAGCAATTTGACCCATCAITGCAATAGCAGGAGTTGTTGAAATGAAATGTTGCTCGGA 2643
Qy 827 PheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeu 846
Db 2644 CAAGCACTGAAGAGTAAGAAAGAACTAGAGGTGCTGGGAAGATCGCTACTGAAGCAATA 2703
Qy 847 GluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGlu 866
Db 2704 GAAACATTCGGAACCGTTGTTCTTTCACTCAGGAGGAGGTTTGAACATATATATGCT 2763
Qy 867 GluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCys 886
Db 2764 CAGAGTTTGAGGTACCATACAGAACTCTTGAGGAAGACACACATCTTTGGAATTACA 2823
Qy 887 TyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAla 906
Db 2824 TTTTCCCTCACCCAGGCAATGATGATATTTTCTCTATGCTGATGTTTCCGCTTTGGAGCC 2883
Qy 907 TyrIleuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIle 926
Db 2884 TACTTTGGGCACATAAACAATCATGAGTCTTGAGGATGTTCTGTAGTATTTTCAGCTGTT 2943
Qy 927 AlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAla 946
Db 2944 GTCTTTGGTGCCATGGCGGGGCAAGTCAGTTTCATTTGCTCTGACTATGCCAAGCC 3003
Qy 947 LysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArg 966
Db 3004 AAATATATCAGCAGCCACATCATCATGATCATTAATGAAAAAACCCTTTGATTCAGAGCTAC 3063
Qy 967 SerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSer 986
Db 3064 AGCAGGAAGGCTTAATGCCGGAACACATTTGGAAGAAATGTTCACATTTGTGTGAAGTTGA 3123
Qy 987 PhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGlu 1006
Db 3124 TTCAACTATCCACCACCGGACATCCAGTGCTTCAGGAGCTGAGCCTGGAGGTGAAG 3183
Qy 1007 ArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGln 1026
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
```

```
Db 3184 AAGGGCCAGCAGCTGCTCTGTTGGGCAGCAGCTGGCTGTGGGAAGACACAGTGGTCCAG 3243
Qy 1027 LeuLeuGlnArgLeuTyrAspProValGlnGlnGlyGlnValLeuPheAspGlyValAspAla 1046
Db 3244 CTCCTGAGCGGTTCTACGACCCCTTGGCAGGAAAGTGTCTGCTGATGGCAAGAAATA 3303
Qy 1047 LysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProVal 1066
Db 3304 AAGCGACTGAATGTTTCAGTGGCTCCGAGCACACCTGGCATCGTGTCCAGGAGCCCATC 3363
Qy 1067 LeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValPro 1086
Db 3364 CTGTTTTCAGTCAGCATTCGTGAGAACATTTGCTATGAGACAAACAGCCGGGTGTCA 3423
Qy 1087 LeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeu 1106
Db 3424 CAGGAAGAGATCGTAGGCGAGCAAGAGGAGCCCAACATACATGCTTCATCGAGTCACTG 3483
Qy 1107 ProGluLysTyrAsnThrGlnValGlyLeuLysGlyValAlaGlnLeuSerGlyGlyGlnLys 1126
Db 3484 CCTAATAAATATAGCACTAAAGTAGGAGACAAAGGAACTCAGCTCTCTGTTGGCCAGAAA 3543
Qy 1127 GlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuAspGlu 1146
Db 3544 CAACGATTTGCCATAGCTCGTGCCTTGTAGACAGCCTCATATTTTGTGTTGGATGAA 3603
Qy 1147 AlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAla 1166
Db 3604 GCCACGTCAGCTCTGATACAGAAAGTGAAGGTTGTCCNAGAAGCCCTGGACAAAGCC 3663
Qy 1167 ArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAsp 1186
Db 3664 AGAGAAAGCGCCACCTGCTATTTGTTGCTCCCGCTGTCCACCATCCAGATGCAGAC 3723
Qy 1187 LeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeu 1206
Db 3724 TTAATAGTGGTGTTCAGAAATGGCAGATGCAGATCAAGGAGCATGGCAGCATCAGCAGTGTG 3783
Qy 1207 ArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
Db 3784 GCACAGAAGGCATCTATTTTCAATGGTCAGTGTCCAGGCT 3825

RESULT 6
US-10-044-671-1
; Sequence 1, Application US/10044671
; Patent No. US20020177147A1
; GENERAL INFORMATION:
; APPLICANT: Washington State University Research Foundation
; APPLICANT: Mealey, Katrina
; APPLICANT: Bentjen, Steven
; TITLE OF INVENTION: MDR1 VARIANTS AND METHODS FOR THEIR USE
; FILE REFERENCE: 4630-61733
; CURRENT APPLICATION NUMBER: US/10/044,671
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/261,578
; PRIOR FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: US 60/314,829
; PRIOR FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 1
; LENGTH: 4317
; TYPE: DNA
; ORGANISM: Canis familiaris
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (70)..(3912)
; OTHER INFORMATION:
US-10-044-671-1

Alignment Scores: 1.51e-313 Length: 4317
Pred. No.: 3525.50 Matches: 680
Score:
```


Percent Similarity: 73.05% Conservative: 239
Best Local Similarity: 54.05% Mismatches: 270
Query Match: 57.24% Indels: 69
DB: 9 Gaps: 9

US-09-873-409-5 (1-1222) x US-10-044-671-1 (1-4317)

```
Qy 1 MetIleLeuGlyIleLeuAlaSerIleuValAenGlyAlaCysLeuProLeuMetProLeu 20
Db 217 ATGTTGGTGGGACAAATGCTGCCATCATCCATGGAGTGCACCTCCCTCTCATGATGCTG 276
Qy 21 ValLeuGlyIleuMetSerAspAsnIleuIleSerGlyCysLeu----- 34
Db 277 GTTTTGGAAACATGACATAGCTTTGCAATGCGAGGAATTTCAAGAAACAAACTTTT 336
Qy 35 -----ValGlnThrAsnThrTyrSerPheArg----- 44
Db 337 CCAGTTATTAATTAAGAAAGTATTACGAACATATACACAACTTTCATCAACCATCTGGAG 396
Qy 45 -----LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe 59
Db 397 GAGGAAATGACCACGATATCCCTATTATTACAGTGGGATCGGTGCTGGTGGCT 456
Qy 60 GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg 79
Db 457 GCTTATCATCCAGGTTTCATCTGCTGCTCGGAGCAGGAGGAGTACTCAAAATTAGA 516
Qy 80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99
Db 517 AAACAAATTTTTCATGCTATCATCGACAGGAGATTGGCTGTGTTGACGTGATGACGTT 576
Qy 100 GlyIleLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
Db 577 GGGAGCTTAACACCGGCTCACAGAGATGCTCCAAATCAATGAAGAAATGCGGAC 636
Qy 119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
Db 637 AAGTTTGAATGTTCTTCAATCAATAGCAACATTTTCCACCGTTTATAGTGGGGTTT 696
Qy 139 ValLysGlyTrpLysLeuValLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
Db 697 ACACCTGGTTGGAAGTAACTACCTTTGATTTTGGCCATCAGCCCTGTTGACCTTCA 756
Qy 159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGlnLeuSerAlaTyrSer 178
Db 757 GCGCCATCTGGGCAAGATACTATCTTCTTACTGATATAAGAACTCTTGCCATGCA 816
Qy 179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198
Db 817 AAGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGAACTGTGATGCTTTGGA 876
Qy 199 AlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPhe 218
Db 877 GCACAAAGAAAGAACTTGAA----- 897
Qy 219 TyrPheProGlnTrpLeuLeuSerCysValLeu**PheValArgTyrThrGlnAsnLeu 238
Db 898 -----AGTACACAAAAATTTA 915
Qy 239 LysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAla 258
Db 916 GAAGAAGCTAAAGAATTTGGATAAAGAAAGCTATCAGGCCCAACATTTCTATTGGTGC 975
Qy 259 ValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIle 278
Db 976 GCTTTCTTATGATCATGATCATATGCTCTGGCTTTCTGGTATGGGACCTCTTTGGTC 1035
Qy 279 LeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIle 298
Db 1036 CTCCTCAGTGAA-----TATACTATGGACAGGTACTCACTGCTCTCTTTCTGATTAT 1089
Qy 299 HisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArg 318
```

```
Db 1090 ATTGGGGCTTTTAGTATTGGACAGCATCCCCAACACANTTGAAGCATTTGCAAAACCAAGA 1149
Qy 319 GlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSer 338
Db 1150 GGCAGAGCTTATGAAATCTTCAAGATAATTGACAATAAACCAAGCATTCACAGCTATTGCG 1209
Qy 339 ThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPhe 358
Db 1210 AAGAGTGACATAAACAGATAATATTAAAGGAAATTTGGAATTCAAAATATTCTCACTTC 1269
Qy 359 AsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSer 378
Db 1270 AGTTACCTCTTCGAAAGAAGATTAAAGATCTTAAGGGTCTCAACTGAAGGTTAGAGT 1329
Qy 379 GlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeu 398
Db 1330 GGCAGACAGTGGCGCTGGTTGGGAACAGTGGCTCGGGAAGAGCAGCACCGCTGCAGCTG 1389
Qy 399 LeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArg 418
Db 1390 ATGCAGAGCTCTATGACCCACAGATGGCATGGTCTGTATTGATGGACAGACATTAGG 1449
Qy 419 AlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeu 438
Db 1450 ACCATAAATGTAAAGGCATCTTCGGGAATTTACTGTGTGGTGGTGCAGGAGCTGTGTTG 1509
Qy 439 PheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspValThrAspGluGlu 458
Db 1510 TTTGCCACCATGATAGCTGAAACCATTCGCTATGCGCGGAAATGTCCACCATGGATGAG 1569
Qy 459 MetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProLeuLys 478
Db 1570 ATTGAGAAAGCTGTTAAAGGAAGCCAAATGCTATGATTTTATCATGAACATACCTAATAA 1629
Qy 479 PheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIle 498
Db 1630 TTTGACACTCTGGTTGGAGAGAGGGGGCCGGCTGAGTGGTGGACAGAAACAGAGAAATC 1689
Qy 499 AlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSer 518
Db 1690 GCATTTGCTGGGCGCTGGTTGCGCAACCCCAAGATTTCTCTGCTGGATGAGGCAAGTCA 1749
Qy 519 AlaLeuAspSerGluSerLysSerAlaValGlnAlaLeuGluLysAlaSerLysGly 538
Db 1750 GCTCTGGACACTGAAAGTGAAGCAGTGGTTTCAGTGGCCCTCGATAAGGCCAGAAAGGC 1809
Qy 539 ArgThrThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleVal 558
Db 1810 CGGACTACCATTTGTATAGCTCATCTGTTGCTACAGTTTCGTAATGCTGATGCTATGCT 1869
Qy 559 ThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArg 578
Db 1870 GGTGTTGATGATGGAGTCAATTTGGAGAAAGGAATCATGATGAACATCATGAAGAGAAG 1929
Qy 579 GlyLeuTyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGlu 598
Db 1930 GGCATTTACTTCAAACTTGTCAATGCGAG---ACAAGAGGAAATGAAATTCAGTTAGAA 1986
Qy 599 SerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSer 618
Db 1987 AATGCCACTGGTGAATCCAAAAAGTGAAGTGAATGCTGCTTGGAAATG---TCTCCAAAGAT 2043
Qy 619 IleLysSerAspPheIle----- 624
Db 2044 TCAGGCTCCAGTTTAAATAAAAGAGATCACTCGCAGAGGATATACATGCACCACCAAGGC 2103
Qy 625 ---AspLysAlaGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeu 643
Db 2104 CAAGACAGAAAGCTTGGTACAAAAGAGGAGCTTGAATGAGAATGCTCTTCAGTTTCTCCTC 2163
Qy 644 LeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAla 663
Db 2164 TGGAGGATTTCTGAAGCTGAACCTCAACTGAATGCGCTTATTTTGTGGTGTGTTATTTGT 2223
```


US-09-873-409-5 (1-1222) x US-09-866-866A-3 (1-3860)

Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
 Db 151 ATGGTGGGAACCTTTGGCTGCCATCATCCATGGGCTGGACTTCCTCTCATGATGCTG 210
 Qy 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
 Db 211 GTGTTTGGAGAAATGACATGATATCTTTGCAAAATGCAAGGAAATTTAGAAATGATGATGCA 270
 Qy 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe 43
 Db 271 AACATCACTAATAGAAAGTATGATATCAATGATACAGGGTTCTTCATGATCTGGAGGAAGAC 330
 Qy 44 -----ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
 Db 331 ATGACCAGATATGCTTATTATACAGTGGAAATTTGGTGGCTGGGTGCTGCTTAC 390
 Qy 62 IleGlnIleSerLeuTrpIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
 Db 391 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAGCAAAATACACAAAATAGAAAACAG 450
 Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
 Db 451 TTTTTCATGCTATAATGGCAGAGGATAGGCTGTTTGTATGTCACCATGTTGGGAG 510
 Qy 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
 Db 511 CTTAACACCCGACTTACAGATGATGCTCTAAGATTAATGAAGTTATTTGGTGCACAAAT 570
 Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
 Db 571 GGAATGTTCTTTCAGTCAATGGCAACATTTTCACTGGGTTTATAGTAGAATTTACACGT 630
 Qy 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
 Db 631 GGTGGAAGCTAACCCCTGTGATTTTGGCCATCAGTCCTGTTCTGGACTGTGAGCTGCT 690
 Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
 Db 691 GTCGGGCAAGATACATCTTCATTTCTGATTAAGAACTCTTAGCGTATGCAAAAGCT 750
 Qy 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
 Db 751 GGAGCAGTAGCTCAAGAGGCTTTGGCAGCAATTAGAACTGTGATGTCATTTGGAGGACAA 810
 Qy 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
 Db 811 AAGAAAGAACTTGAA----- 825
 Qy 221 ProGlnTrpLeuLeuSerCysValLeu**PheValArgTyrThrGlnAsnLeuLysAsp 240
 Db 826 -----AGTACAAACAAAATTTAGAGAA 849
 Qy 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
 Db 850 GCTAAAGAAATTTGGGATAAGAAAGCTATTACAGCCAAATTTCTATAGTGCTGCTTTC 909
 Qy 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
 Db 910 CTGCTGATCTATGCATCTATGCTCTGGCCCTTCTGGTATGGGACCACTTGGTCTCTCA 969
 Qy 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
 Db 970 GGGGAA-----TATTCTATTGGCAAGTACTCACTGATTTCTTTCTGATTAATTTGGG 1023
 Qy 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
 Db 1024 GCTTTTAGTGTGGACAGGATCTCCAGAGCTTGAAGCATTTGCCAAATGCAAGAGGCA 1083
 Qy 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
 Db 1084 GCTTATGAAATCTTCAAGATAATTTGATAATAAGCCAAAGTATTGACAGCTATTTCGAAGAGT 1143

Qy 341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
 Db 1144 GGGCACAACACGATAATATTAAAGGAAATTTGGAATTCAGAAATGTTTCACATTCAGTTAC 1203
 Qy 361 ProSerArgProSerIleLysIleLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
 Db 1204 CCATCTCGAAAAGATTTAAGATCTTGAAGGCGCTGAACCTGAAGGTGCAGAGTGGGAG 1263
 Qy 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
 Db 1264 ACGETGGCCCTGTTGGAAACAGTGGCTGTGGAGAGACACACAGTCCAGCTGATGTCAG 1323
 Qy 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
 Db 1324 AGGCTCTATGACCCACAGAGGGGATGTCAGTGTGTGATGGACAGGATATTAGGACCAT 1383
 Qy 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
 Db 1384 AATGTAAGTGTTCACGGGAAATCATTTGGTGGTGGTGCAGAACTGTATTTGTTGCC 1443
 Qy 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
 Db 1444 ACCAGTAGCTGAAACATTCGCTATGCGCGTGAAATGTCCACATGATGAGATTGAG 1503
 Qy 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
 Db 1504 AAGCTGTCAAGGAAGCCAAATGCTATGATCTTATCATGAAACTGCTCATAAATTTGAC 1563
 Qy 481 ThrLeuValGlyGlyLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
 Db 1564 ACCCTGTTGGAGAGAGGGGCCAGTTGAGTGGTGGGCAAGCAGAGGATTCGCCATT 1623
 Qy 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
 Db 1624 GCACGTGCCCTGGTTCGCAACCCCAAGATCTCTGCTGATGAGGCCACGTCAGCTTGG 1683
 Qy 521 AspSerGluSerLysSerAlaValGlnAlaLeuGluLysAlaSerLysGlyArgThr 540
 Db 1684 GACACAGAAAGCGAAGCAGTGGTTGAGTGGTGGGCAAGCAGAAAGGTCCGACC 1743
 Qy 541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560
 Db 1744 ACCATTTGATAGCTCATCGTTTGTCTACAGTTGCTGCTGATGAGGCCACGTCAGCTTTC 1803
 Qy 561 LysAspGlyMetLeuAlaGlyLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580
 Db 1804 GATGATGGAGTCTATGTGGAGAAAGAAATCATGATGAATCATGAAAGAAAGGCATT 1863
 Qy 581 TyrTyrSerLeuValMetSerGln-----AspIleLysLysAlaAsp 594
 Db 1864 TACTTCAAACTTGTCAATGCGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCT 1923
 Qy 595 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis 614
 Db 1924 GATGATCCAAAAGTGAATTTGATGCTTTGGAAATGCTCTTCAATGATTCAGATCCAGT 1983
 Qy 615 SerVal-----LysSerIleLys---SerAspPheIleAspLysAla 627
 Db 1984 CTAATAAGAAAAGATCAACTCGTAGGAGTGTCCGTGGATGCACAGCCCAAGACAGAAAG 2043
 Qy 628 GluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeu 647
 Db 2044 CTTAGTACCAGAGGCTCTGGATGAAGATATACCTCCAGTTTCTTTTGGAGGATTATG 2103
 Qy 648 LysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsn 667
 Db 2104 AAGCTAAATTTAACTGAATGGCTTATTTTGTGTGTGTTTGTGCTATTTTGTGCCATTAAAT 2163
 Qy 668 GlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn 687
 Db 2164 GGAGGCTTCAACACGACATTTGCAATAATATTTCGAAGATTATTAGGGGTTTTTACAAGA 2223

Qy	688	----	AsnAspLysThrThrLeuLysHisAspAlaGluLeuTyrSerMetIlePheValIle	706
Db	2224	ATTGATGATCCTCAACAAAAACGACAGAAATAGTAAC	TTTTTTTTCACATATGTTTCTTCAGCC	2283
Qy	707	LeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyAAspAlaGly	726	
Db	2284	CTTGGAAATATTCTTTTATACATTTTTCTTCAGGTTTCACATTTGCCAAGCTGGA	2343	
Qy	727	GluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIle	746	
Db	2344	GAGATCCTCACCAAGCGGCTCCGATACATGTTTCCGATCCATGCTCAGACAGGATGTG	2403	
Qy	747	AlaTrpPheAspGlnLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAsp	766	
Db	2404	AGTTGGTTGATGATGCCCTAAACACACCACTGGAGCATTTGACTACCAAGCTCGCAATGAT	2463	
Qy	767	IleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThr	786	
Db	2464	GCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTGCTGTAAATTATCCCAAGAAATAGCA	2523	
Qy	787	AsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIle	806	
Db	2524	AATCTTGGGACAGGAATATTATATCTTCATCTATGTTGGCAACTAAACACTGTTTACTC	2583	
Qy	807	LeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGly	826	
Db	2584	TTAGCAATTGTACCCATCATTCGAAATAGCAGGAGTTGTGAATATGAAAATGTTGCTCTGGA	2643	
Qy	827	PheAlaAsnLysAspLysGlnGluIleuLysHisAlaGlyLysIleAlaThrGluAlaLeu	846	
Db	2644	CAAGCACTGAAGATGAAGAAAGCACTAGAAAGGTGCTGGGAAGATCGCTACTCAAGCAATA	2703	
Qy	847	GluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGlu	866	
Db	2704	GAATACTTCGAACCGGTTGTTCTTTGACTCAGGAGCAGAGTTTGAACATATGTTATGCT	2763	
Qy	867	GluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCys	886	
Db	2764	CAGAGTTTGCAGGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTTGGAATATCA	2823	
Qy	887	TyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAla	906	
Db	2824	TTTTCTTCACCCAGGCAATGATGATTTTTCTTATGCTGATGTTTCCGGTTTGGAGCC	2883	
Qy	907	TyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIle	926	
Db	2884	TACTTGTGGCACATAAACTCATGAGCTTTGAGGATGTTCTGTAGTATATTTTCAGCTGTT	2943	
Qy	927	AlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAla	946	
Db	2944	GTCTTTGGTCCCATGGCCGTGGGGCAAGTCAGTTTCATTTGCTCCTGACTATGCCAAAGCC	3003	
Qy	947	LysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArg	966	
Db	3004	AAATATACAGCAGCCACATCATCATGATCATTTGAAAAAACCCCTTTGATTGACAGCTAC	3063	
Qy	967	SerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSer	986	
Db	3064	AGCAGGAAGGCTTAATGCCGAACATTTGGAAGGAATGTCACATTTGGTGAAGTTGTA	3123	
Qy	987	PhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGlu	1006	
Db	3124	TTCAACTATCCCACCGACCGGACATCCCAGTGCTTCAGGACTGAGGCTGAGGTTGAAG	3183	
Qy	1007	ArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGln	1026	
Db	3184	AAGGCCAGACGCTGGCTCTCGTGGGCAGCAGCTGGCTGTGGGAAGAGACACAGTGGTCCAG	3243	
Qy	1027	LeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAla	1046	
Db	3244	CTCTCGAGCGGTTCTACGACCCCTTTGGCAGCGGAAGTGTGCTTGTGGAAGAAATAA	3303	
Qy	1047	LysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProVal	1066	

```

Db 3304 AGCGCATGATGTTTCAGTGGCTCCGAGCACACCTGGGCATCGTCCCGAGGCCCCATC 3363
Qy LeuPheAsnCyssSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValPro 1086
Db 3364 CTGTTTGACTGCAGCATTCCTGAGAACATTCCTATGAGACAAACAGCCGGTGTGTCA 3423
Qy LeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeu 1106
Db 3424 CAGGAAGAGATCGTGAGGGCAGCAAGGAGGCCAACATACATGCTTCATCGAGTCACTG 3483
Qy ProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLys 1126
Db 3484 CCTAATAATATAGCACTAAGATGAGAGCAAGAGACTCAGCTCTCTGTGTGCCAGAAA 3543
Qy GlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGlu 1146
Db 3544 CAACGCATTGCCATAGCTCGTGCCCTTGTTAGACAGCCTCATATTTCCTTTGGATGAA 3603
Qy AlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAla 1166
Db 3604 GCCACGTGAGCTCTGGATACAGAAAGTGAAGAGTTGTCCAAGAAAGCCCTGGACAAAGCC 3663
Qy ArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAsp 1186
Db 3664 AGAAGAGCCGCACTGATTCATGTTGTCACCGCCTGTCCACCATCCAGANTGCAGAC 3723
Qy LeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlnLeuLeu 1206
Db 3724 TTATAGTGGTGTTTCAGATGCGCAGATGCGAGGAGCATGGCAGCATCAGCAGCTGCTG 3783
Qy ArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
Db 3784 GCACAGAAAGGCATCTATTTTCAATGGTCAGTGTCCAGGCT 3825

RESULT 8
US-09-306-417-1
; Sequence 1, Application US/09306417
; Patent No. US20020103144A1
; GENERAL INFORMATION:
; APPLICANT: Heinrich-Pette-Institut
; TITLE OF INVENTION: Retroviral Gene Transfer Vectors
; FILE REFERENCE: P50491
; CURRENT APPLICATION NUMBER: US/09/306,417
; CURRENT FILING DATE: 1999-05-06
; EARLIER APPLICATION NUMBER: DE 198 22 115
; EARLIER FILING DATE: 1998-05-08
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 1
; LENGTH: 8630
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: proviral
; OTHER INFORMATION: plasmid DNA
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(160)
; OTHER INFORMATION: plasmid backbone (pUC)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (161)..(677)
; OTHER INFORMATION: 5'-LTR
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (532)..(1219)
; FEATURE:
; NAME/KEY: mat peptide
; LOCATION: (1220)..(5062)
; OTHER INFORMATION: m4 mdr-1 cDNA
; FEATURE:
; NAME/KEY: misc feature

```


; LOCATION: (5215)..(5774)
; OTHER INFORMATION: 3'-LTR
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (5775)..(8630)
; OTHER INFORMATION: plasmid backbone (pUC)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(8630)
; OTHER INFORMATION: retroviral expression vector SFbeta71m4
US-09-306-417-1

Alignment Scores:
Pred. No.: 9.87e-313 Length: 8630
Score: 3521.50 Matches: 681
Percent Similarity: 73.92% Conservative: 246
Best Local Similarity: 54.31% Mismatches: 264
Query Match: 57.18% Indels: 63
DB: 10 Gaps: 10

US-09-873-409-5 (1-1222) x US-09-306-417-1 (1-8630)

QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
DB 1370 ATGGTGGTGGAACTTTGGCTGCCATCATCCATCGGGCTGGACTTCTCTCATGATGCTG 1429
QY 21 ValLeuGlyGluMetSerAsp-----AenLeuIleSer 31
DB 1430 GTGTTGGAGAAATGACAGATATCTTTGCCAAATGCAGGAATTTAGNAGATCTGATGTCA 1489
QY 32 GlyCysLeuValGlnThr-----AenThrTySerPhePhe----- 43
DB 1490 AACATCACTAATAGAGTGCATATCATGATACAGGGTCTTCTCATGATCTGGAGGAGAC 1549
QY 44 -----ArgLeuThrLeuTyThrValGlyIleGlyValAlaAlaLeuIlePheGlyTy 61
DB 1550 ATGACCATGATGCCATTATTATACAGTGAATTTGGTCTGGGTGCTGGTCTGCTTAC 1609
QY 62 IleGlnIleSerLeuTrpIleThrAlaAlaArgGlnThrIleArgIleArgGlyGln 81
DB 1610 ATTCAGTTTTCATTTTGGTCCCTGGCAGCTGGAGACAAATACACAAATTAGAACACAG 1669
QY 82 PhePheHisValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
DB 1670 TTTTTCATGCTAATAATCGCAGCAGAGATAGGCTGGTTGATGTCACGATGTTGGGAG 1729
QY 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
DB 1730 CTTAACACCCGACTTACAGATGATGCTCTTAAGATTAAATGAAGTTATTGGTGACAAAT 1789
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
DB 1790 GGAATGTTCTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGATTTACACGT 1849
QY 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
DB 1850 GGTGGAAGCTAACCTTGATTTTGGCCATCAGTCTCTTCTGGAGCTCAGCTGCT 1909
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTySerLysAla 180
DB 1910 GTCTGGGCAAGATATATCTTCTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 1969
QY 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
DB 1970 GGAGCAGTAGCTGAAGAGTCTTGGCAGCAATTTAGAACTGTGATTCATTTGGAGACAA 2029
QY 201 GlnLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyTrpAlaTrpPheTyPhe 220
DB 2030 AAGAAAGAACTTCAA----- 2044
QY 221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyThrGlnAsnLeuLysAsp 240
DB 2045 -----AGGTACAAACAAATTTAGAAAG 2068

QY 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTy 260
DB 2069 GCTAAAGAAATTGGGATAAAGAAAGCTATTACAGCAATATTCTATAGTGCTGCTTTC 2128
QY 261 PhePheMetAsnGlyThrTyThrGlyLeuAlaPheTyThrGlyThrSerLeuIleLeuAsn 280
DB 2129 CTGCTGATCTATGCACTTATGCTCTGGCCCTCTGTGTGGGACCACTTGGTCTCTCA 2188
QY 281 GlyGluProGlyTyThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
DB 2189 GGGGAA-----TATCTATTGGACAAGTACTCCTGTTATTTCTTCTGTTAATTTGG 2242
QY 301 SerTyrcysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
DB 2243 GCTTTTATGTTGGACAGGCATCTCCAGCATTTGAAGCATTTGCAAAATGTCAGAGGAGCA 2302
QY 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
DB 2303 GCTTATGAAATCTTCAAGATAATTGATAATAGCCCAAGTATTGCACAGCTATTGCAAGAGT 2362
QY 341 GlyTyrcysProGluSerIleGlyThrValGluPheLeuAsnValSerPheAsnTy 360
DB 2363 GGGCACAACCAAGATAATATTAAAGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 2422
QY 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
DB 2423 CCATCTCGAAAAAGAGTTAAGATCTTGAAGGGCTTGAACCTGAAGGTGCAGAGTGGCAG 2482
QY 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
DB 2483 ACGGTGGCCCTGGTTGGAAACAGTGGCTGTGGAGAGACACACAGTCCAGCTGATGTCAG 2542
QY 401 ArgLeuTyrcysProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
DB 2543 AGGCTCTATGACCCACAGAGGGGATGGTCAGTGTGTGATGGACAGATATTAGACCATA 2602
QY 421 AsnValArgHisTyrcysArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
DB 2603 AATGTAAGGTTTCTACGGGAAATCATTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2662
QY 441 ThrThrIleSerAsnAsnIleLysTyrcysArgAspAspValThrAspGluGluMetGlu 460
DB 2663 ACCACGATGATGAAACATTCGCTATGCGCCGTGAAATGTCCACCATGATGATGATGATG 2722
QY 461 ArgAlaAlaArgGluAlaAsnAlaTyrcysPheIleMetGluPheProAsnLysPheAsn 480
DB 2723 AAAGCTGTCAAGAAAGCCAAATGCTATGACTTTATCATGAAACTGCTCATAAATTTGAC 2782
QY 481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
DB 2783 ACCCTGTTGGAGAGAGAGGGCCAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2842
QY 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
DB 2843 GCACGTGCCCTGGTTCGCAACCCCAAGATCCTCTCTGGATGGAGCCAGCAGTCCAGCTTG 2902
QY 521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThr 540
DB 2903 GACACAGAAGCGAAGCAGTGGTTCAGGTGGTCTCGATGAAGCCAGAAAGTCCGAGCC 2962
QY 541 ThrIleValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560
DB 2963 ACCATTGTGATAGTCTATCGTTTGTCTACAGTTCTGTAATGCTGACGCTCATCGTGTTC 3022
QY 561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580
DB 3023 GATGATGGAGTCTATTGTGGAGAAAGGAAATCATGATGAATCATGATGAAGAGAAAGCAT 3082
QY 581 TyrcysSerLeuValMetSerGln-----AspIleLysLysAlaAsp 594
DB 3083 TACTTCAAACTTGTCAATGCAGACAGCAGGAAATGAAGTTGAATTAGAAATTCAGCT 3142

FEATURE: OTHER INFORMATION: Description of Artificial Sequence: proviral
OTHER INFORMATION: plasmid DNA
FEATURE: NAME/KEY: misc feature
LOCATION: (1)..(8630)
OTHER INFORMATION: retroviral expression vector SFbeta91msA1
FEATURE: NAME/KEY: misc feature
LOCATION: (1)..(160)
OTHER INFORMATION: plasmid backbone (pUC)
FEATURE: NAME/KEY: misc feature
LOCATION: (161)..(677)
OTHER INFORMATION: 5'-LTR
FEATURE: NAME/KEY: 5'UTR
LOCATION: (532)..(1219)
FEATURE: NAME/KEY: mat_peptide
LOCATION: (1220)..(5062)
OTHER INFORMATION: msA1 mdr1 cDNA
FEATURE: NAME/KEY: misc feature
LOCATION: (5215)..(5774)
OTHER INFORMATION: 3'-LTR
FEATURE: NAME/KEY: misc feature
LOCATION: (5775)..(8630)
OTHER INFORMATION: plasmid backbone (pUC)
US-09-306-417-2

Alignment Scores:
Pred. No.: 9,87e-313 Length: 8630
Score: 3521.50 Matches: 681
Percent Similarity: 73.92% Conservative: 246
Best Local Similarity: 54.31% Mismatches: 264
Query Match: 57.18% Indels: 63
DB: 10 Gaps: 10

US-09-873-409-5 (1-1222) x US-09-306-417-2 (1-8630)

QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 1370 ATGGTGGTGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 1429
QY 21 ValLeuGlyGluMetSerAsp-----AnLeuIleSer 31
Db 1430 GTGTTTGGAGAAATGACAGATATCTTTGCAAAATGCAGAAATTTAGAAATCTCATGATGCA 1489
QY 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
Db 1490 AACATCACTAATAGAGTATCAATGATACAGGGTTCCTTGAATCTGGAGGAGAC 1549
QY 44 -----ArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db 1550 ATGACCAAGGTATGCTTATTATACAGTGAATGGTCTCGGGTGGTGGTGGCTTAC 1609
QY 62 IleGlnIleSerLeuThrIleLeuAlaAlaArgGlnThrIleArgIleGlyGln 81
Db 1610 ATTCAGGTTTCATTTGGTGGCTGGCAGCTGGAAGACAAATATACAAAAATTAGAAAAACAG 1669
QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db 1670 TTTTTCATGCTATAATGGCAGGAGATAGGCTGGTTTATGATGTCACGATGTTGGGGAG 1729
QY 102 LeuAsnThrArgMetThr-----AspIleAspIleSerAspGlyIleGlyAspLysile 120
Db 1730 CTTAAACCCGACTTACAGATGATGCTCTTAAGATTAATGAATTTATGTTGTTGTTGTTG 1789
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db 1790 GGAATGTTCTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGGATTTACACGT 1849

QY 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db 1850 GGTGGAGACTAAACCTTGATTTGGCCATCAGTCTCTGTTCTTGGACTGTGCTGCT 1909
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db 1910 GTCGGGCAAGATATCTATCTTACTGATAAAGAACTCTTACGCTATGCAAAAGCT 1969
QY 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db 1970 GGAGCAGTAGCTGAAGAGCTTTGGCAGCAATTAGAATGTGATTCATTTGGAGGACAA 2029
QY 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
Db 2030 AAGAAAGAACTTGAA----- 2044
QY 221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
Db 2045 -----AGGTACAAACAAAATTTAGAGAA 2068
QY 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
Db 2069 GCTAAAGAAATTTGGATAAAGAAAGCTATTACAGCCAATATTCTATAGTGTCTGCTTC 2128
QY 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
Db 2129 CTGCTGATCTATGATCTTTATGCTCTGGCTTCTGTTATGGGACCACTTGTGCTCTCA 2188
QY 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
Db 2189 GGGGAA-----TATTCTATTGGACAAAGTACTCACTGTATTCTTTCTGTATTAAATGGG 2242
QY 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
Db 2243 GCCTTTAGTGTGGACAGCACTCCAGCATTTGAAGCAATTTGCAATGCAAGAGAGCA 2302
QY 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
Db 2303 GCTTATGAAATCTTCAAGATAATTGATAATAGCCAAATATTGACAGCTATTGCAAGAGT 2362
QY 341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
Db 2363 GGGCACAAACCAATATATTAAAGGAAATTTGGAATTTGCAAAATGTTCATCTCAGTTAC 2422
QY 361 ProSerArgProSerIleLysIleLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
Db 2423 CCATCTCGAAAGAAAGTTAAGATCTTGAAGGCTGAACTTGAAGTGCAGATGGGCGAG 2482
QY 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
Db 2483 ACGTGGCCCTGTTGGAAACAGTGGCTGGGAGAGACACACAGTCCAGCTGATGTCAG 2542
QY 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
Db 2543 AGGCTCTATGACCCACAGAGGGGATGGTCAGTGTGTGATGGACAGGATATTAGACCATTA 2602
QY 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
Db 2603 AATGTAAGGTTTCTACGGGAAATCATTTGGTGTGTGAGTCAAGAACTGTATTGTTGCC 2662
QY 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspValThrAspGluMetGlu 460
Db 2663 ACCACGATAGCTGAAACCATTCGCTATGGCCGTGAAATGTCCACCTGATGATGATGAG 2722
QY 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
Db 2723 AAAGTGTCAAGGAAGCCAAATGCTATGACTTTATCATGAACTGCGCTCATATAAATTGAC 2782
QY 481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
Db 2783 ACCCTGGTTGGAGAGAGGGGCCCAAGTTAGTGGTGGGAGAGAGGATGCGCCATT 2842

Db 5003 GCACAGAAAGGCATCTATTTTCAATGGTCAGTGTCCAGGCT 5044

RESULT 10

US-09-769-097-3

; Sequence 3, Application US/09769097

; Patent No. US20020055128A1

; GENERAL INFORMATION:

; APPLICANT: Kimberly Anne Brun

; APPLICANT: Richard James Chenery

; APPLICANT: Harma Ellens

; APPLICANT: John Anthony Feild

; APPLICANT: Lin Yue

; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES

; TITLE OF INVENTION: ENCODING RAT MORIA AND SCREENING METHODS THEREOF

; FILE REFERENCE: GP-5009-C2

; CURRENT APPLICATION NUMBER: US/09/769,097

; PRIOR FILING DATE: 2001-01-25

; PRIOR APPLICATION NUMBER: 09/208,809

; PRIOR FILING DATE: 1998-12-09

; PRIOR APPLICATION NUMBER: 09/156,800

; PRIOR FILING DATE: 1999-09-17

; PRIOR APPLICATION NUMBER: US99/20770

; PRIOR FILING DATE: 1999-09-10

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 3

; LENGTH: 4425

; TYPE: DNA

; ORGANISM: RATTUS RATTUS

US-09-769-097-3

Alignment Scores:

Pred. No.:	9,44e-313	Length:	4425
Score:	3517.00	Matches:	680
Percent Similarity:	73.90%	Conservative:	243
Best Local Similarity:	54.44%	Mismatches:	266
Query Match:	57.10%	Indels:	60
DB:	10	Gaps:	10

US-09-873-409-5 (1-1222) x US-09-769-097-3 (1-4425)

Qy	1	MetIleuGlyIleuLeuSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu	20
Db	499	ATGCTGTGGAACTCTGGCGGCATTTCCATGGAATTCGCTCCCACTTATGATGCTG	558
Qy	21	ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr	40
Db	559	GTCTTTGGAGACATGACATAGCTTTTGCAAAAT-----GTAGGAAACAAACCGTAGTATG	612
Qy	41	SerPhePheArgLeuThr-----	46
Db	613	AGTTTCTACAATGCTACAGACATATATGCCAAGCTGGAGGACGAAATGACCACGTAGCC	672
Qy	47	LeuTyrTyrValGlyIleGlyValAlaLeuIlePheGlyTyrIleGlnIleSerLeu	66
Db	673	TACTATTACAGGCGCATTTGGTGGCGGTGCTCATCGTTGCCATACATCCAGGTTTCCACT	732
Qy	67	TripIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGlnPheHisSerVal	86
Db	733	TGTGTGCTGCAGCTGGGACAAATACACAGATAGGCAGAGTTTTTCCATGCCATC	792
Qy	87	LeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGluLeuAsnThrArgMet	106
Db	793	ATCAATCAGAGATAGGCTGGTTTGAGTGCATGACGCTGGGGAGCTCAACACCGGCTC	852
Qy	107	Thr---AspIleAspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGln	125
Db	853	ACAGATGACGTCTCRAAAATTAATGAAGGAATGGTGACAAAATGGAAATGTTCTTTTCAG	912
Qy	126	AsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLysGlyTyrLysLeuThr	145
Db	913	GCAATGCAACATTTTTTGTGTGTTTTTATAATAGGATTTTACTCGCGCTCGGAAGCTAACT	972

Qy	146	LeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAlaCysSerArgMet	165
Db	973	CTTGTGATTTGGCCATCAGCCCTGTTCTTGACTGTCTGAAGATA	1032
Qy	166	ValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGlu	185
Db	1033	TTGTCTTCAATTTACTGATAAGAACTCCAGGCTTATGCAAAAGCTGGAGCAGTTGCTGAA	1092
Qy	186	GluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGlnLysGluLeuGln	205
Db	1093	GAAGTCTTTAGCAGCCATCAGAACTGTGATTGCTTTGGAGGACAAAAGAAAGGAACCTTGAA	1152
Qy	206	ArgSerPheLeuLeuAsnIleThrArgTyrAlaTyrPheTyrPheProGlnTrpLeuLeu	225
Db	1152	-----	1152
Qy	226	SerCysValLeu**PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly	245
Db	1153	-----AGGTACAATAACAATTTGGAAGAGCTAAAGGCTTGGG	1191
Qy	246	IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGly	265
Db	1192	ATAAGAAAGCTATCAGCGCCAACTTCATGGGTGACGCTTTCTGCTTATCTATGCA	1251
Qy	266	ThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr	285
Db	1252	TCATATGCTCTGGCATTCGGTATGGGACTTCCTTGGTCATCTCAAAAGAA-----TAC	1305
Qy	286	ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly	305
Db	1306	ACTATTGGACAAGTGTCTACTGCTTTTTTTTCTGTATTATTTGGAGCATTCAGTGTGGG	1365
Qy	306	AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe	325
Db	1366	CAGGCATCTCCAAATATTGAAGCTTTCGCAATGCTTAGAGGAGCAGCTTATGAAGCTTC	1425
Qy	326	GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu	345
Db	1426	AGTATAATTTGATAAATAAGCCCACTAGTATAGACAGCTTCTCAAGAGAGTGGGCACAAACCCGAC	1485
Qy	346	SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer	365
Db	1486	ACATACAAGAAATTTGGAATTCAAAATATTCACTTCAGTTACCCCTCGAAAGAC	1545
Qy	366	IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal	385
Db	1546	GTCAGATCTTGAAGGCTCAACCTGAAGGTCAAGAGCGGCGCAGACGCTAGCCCTGGTT	1605
Qy	386	GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuGlnArgLeuTyrAspPro	405
Db	1606	GGCAACAGTGGCTGTGGGAAAAGCACAACTGTCAGCTGTCAGAGGCTCTACGACCCC	1665
Qy	406	AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr	425
Db	1666	ATAGAGGCGAGTCAGTATCGACGCGACAGGACATCAGACCACTCAATGTGAGGTATCTG	1725
Qy	426	ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn	445
Db	1726	CGGAAATCATTTGGGTGTGAGTCAGGAACCCGTCGTTTCCACCACCAATTCGCGAA	1785
Qy	446	AsnIleLysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGlu	465
Db	1786	AACATTCGTCATGGCGGAGAAAACGTCACATGGATGATAGAGAAAGCTGTCACAGGAA	1845
Qy	466	AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu	485
Db	1846	GCCAAATGCTTATGATTCATGATAAACTGCCCAAAATTTGACACCTTGGTGTGGTGG	1905
Qy	486	LysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal	505
Db	1906	AGAGGGCGGCGAGCTGATGGGGGACAAAACAGAGATGCCATTCGCCCGGCGCTGCTC	1965
Qy	506	ArgAsnProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys	525

Db 1966 CGAACCCCAAGATCCTTTTGGATGAGCCACGTCAGCCTTGGACACAGAAAGCGAA 2025
Qy 526 SerAlaValGlnAlaAlaLeuGluIysAlaSerIysGlyArgThrThrIleValValAla 545
Db 2026 GCCGTGGTTCAGGCCGCTCTGGATAGAGCTAGAGAGCGGACCCACCATTTGTGATGCT 2085
Qy 546 HisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuIysAspGlyMetLeu 565
Db 2086 CACCGCTTGTCTACAGTTCGCAATGCTGATGTCATCTGCTGGTTTGTGATGGTGTCAAT 2145
Qy 566 AlaGluIysGlyAlaHisAlaGluLeuMetAlaIysArgGlyLeuIleThrVal 585
Db 2146 GTGAGCAAGGAATCATGATGACTCATGAGAGAGAAAGAAATTAATCTTCAAACTTGTC 2205
Qy 586 MetSerGln-----AspIleIysAlaAspGluGlnMetGluSer----- 599
Db 2206 ATGACTCAGACAGCAGGAATGAATTAATGAATAGGAATGAAGCTTGTGAATCTAAAGAY 2265
Qy 600 -----MetThrTyrSerThrGluArgIysThrAsnSerLeu-----ProLeu 613
Db 2266 GGAATTGATATGTGGACATGCTTCAAGAGATTCRGATCCAGTCTAATAAGNAGAGA 2325
Qy 614 HisSerValIysSerIleLysSerAspPheIleAspIysAlaGluGluSerThr---Gln 632
Db 2326 TCAACTCGCAAAAGCATCGTGGGCCACATGATCAAGACGGGGAATTAGCACCAGAGAG 2385
Qy 633 SerIysGluIleSerLeuProGluValSerLeuLeuIysIleLeuIysLeuAsnLysPro 652
Db 2386 GCTCTGATGACGACGACCTCCAGCTCCCTTTGGCGGATCCCTGGAATGGAATCAACT 2445
Qy 653 GluTrpProPheValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPro 672
Db 2446 GAATGGCTTATTTTGGTGGTGTATTTGTGCCATATAAATAGAGGCTTGCAACCA 2505
Qy 673 ValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThr 692
Db 2506 GCAATCTCCATAATATTCAAGAGTGTAGGGGTTTTTCAAAAAAATGACACCCCTGAA 2565
Qy 693 Leu---LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCys 711
Db 2566 ATCCAGCGGCAAGACAGCACTTCTTTTATTTATTTCTGATCCCTGGATCATCTCT 2625
Qy 712 PheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyIleLeuThrMet 731
Db 2626 TTCAATAGCTTTTCTTCAGGGCTTCACTTGGCAAGCTGGAGAGATCTTCAACCAAG 2685
Qy 732 ArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGlu 751
Db 2686 CGACTCCGATACATGCTTCAAAATCCATGCTGAGACAGGACATAAGCTGGTTGATGAC 2745
Qy 752 LysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGln 771
Db 2746 CCTAAACACACAGGAGCGCTGACACACAGGGCTTGGCAATGACGCTGCTCAAGTAAA 2805
Qy 772 GlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSer 791
Db 2806 GGGCTACAGGGCTTAGGCTTGTATTATCCAGAACATAGCAATCTTGGACAGGC 2865
Qy 792 ValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaPro 811
Db 2866 ATCATCATATCCCTGATCTACGGCTGGCAATGACATTTTACTCTTAGCAATTTGTTCCC 2925
Qy 812 ValLeuAlaValThrGlyMetIleGluThrAlaIleMetThrGlyPheAlaAsnLysAsp 831
Db 2926 ATCATGTATAGCAGGAGTGGTTGAAATGAAATGTTGTCTGCAAGCGCTGAAAGAT 2985
Qy 832 LysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThr 851
Db 2986 AAGAAGGAACCTAGAGGTTCTGGGAGAGATCCCTACAGAGGCAATGAAATTTTCGCACT 3045
Qy 852 IleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThr 871
Db 3046 GTCGTCTCTTTGACTCGGAGCAGAAAGTTTGAAACTATGTATGATGCCAGAGCTTGCAGATA 3105

Qy 872 GlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHis 891
Db 3106 CCATACAGAAATGCTTTGAGAAAGCGCAGTCTTTGGGATCACTTTCTCTTCAACCAG 3165
Qy 892 AlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAla 911
Db 3166 GCCATGATGATTTCTCTTATGCTGCTGCTTTCGCGTTTGTATGCTTACTTGGTGGCAGCA 3225
Qy 912 GlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMet 931
Db 3226 GAATCATGACATTTGAAATGTTCTGTAGTATTTCTAGCTATTTGTTGTGGCATG 3285
Qy 932 AlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAla 951
Db 3286 GCAGTGGGAGGTCACTTCATTCGCTCTCTGACTACGCGAAGCCAAAGTCTCGGCATCC 3345
Qy 952 HisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLys 971
Db 3346 CACATCATCAGGATCATTTGAGAAATCCCTGAGATTTGACAGCTTACAGCAGCGAGGCTTG 3405
Qy 972 LysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCys 991
Db 3406 AAGCCTAATATGTTGAAAGGAAATGTAATTAATGGAGTCATGTTCAACTATCCACC 3465
Qy 992 ArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrVal 1011
Db 3466 CGACCCCAACATCCAGTGTCTCAGGGGCTGAGCTAGAGGTGAAGAAAGGCGACGCTG 3525
Qy 1012 AlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeu 1031
Db 3526 GCCTCGTGGGACAGTGGCTCGGGAAGAGTACAGTGGTCCAGCTGCTTGGCGCTTC 3585
Qy 1032 TyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnVal 1051
Db 3586 TATGACCCCATGCCGGAACAGAGTGTTCCTAGATGGCAAAATAAGAACCACTCAATGTC 3645
Qy 1052 GluTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSer 1071
Db 3646 CAGTGGCTCGCGCCACCTGGGCAATGTGTCAGGAGGCCATCTCTGTTGACTGAGC 3705
Qy 1072 IleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLys 1091
Db 3706 ATCCCGCAGAACTTGCCTACGGAGACAACAGCGGTGCTGCTCTCATAAGGAGATCGTG 3765
Qy 1092 GluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsn 1111
Db 3766 AAGGCAAGCCAGGAGGCAACATCCACAGTTCATCGACTCCTGCTGAGAAATACAC 3825
Qy 1112 ThrGlnValGlyLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIle 1131
Db 3826 ACAGAGTGGAGACAAGGACTCAGTCTCGGCGGGCAGAGCAGCGCATCGCCATC 3885
Qy 1132 AlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeu 1151
Db 3886 GCGCGGCCCTCGTCAGACAGCCTCATCTTACTTCTGATGAAGCGACATCAGCTCTG 3945
Qy 1152 AspAsnAspSerGluIysValValGlnHisAlaLeuAspLysAlaArgThrGlyValThr 1171
Db 3946 GATACGAGAGTGAAGGTCGTCGAGAGAGCGCTGGACAAAGCCAGGAGGCGCGCAC 4005
Qy 1172 CysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeu 1191
Db 4006 TGCATTTGTATCGGCACCCGCTGTCCCATCCAGAACCGCAGACTTGTATCGTGTGATT 4065
Qy 1192 HisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIle 1211
Db 4066 CAGAAACGGCAGGTCAAGGAGCAGCGCACCCACAGCAGCTGCTGGCCCGAGAAAGGCATC 4125
Qy 1212 TyrPheLysLeuValAsnAlaGlnSer 1220
Db 4126 TATTTCTCGATGGTCACTGTGCAGGCT 4152

RESULT 11

US-09-769-097-1
; Sequence 1, Application US/09769097
; Patent No. US20020055128A1
; GENERAL INFORMATION:
; APPLICANT: Kimberly Anne Brun
; APPLICANT: Richard James Chenery
; APPLICANT: Harma Ellens
; APPLICANT: John Anthony Feild
; APPLICANT: Lin Yue
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES
; TITLE OF INVENTION: ENCODING RAT MDRIA AND SCREENING METHODS THEREOF
; FILE REFERENCE: GP-50009-C2
; CURRENT APPLICATION NUMBER: US/09/769,097
; CURRENT FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 09/208,809
; PRIOR FILING DATE: 1998-12-09
; PRIOR APPLICATION NUMBER: 09/156,800
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: US99/20770
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 4369
; TYPE: DNA
; ORGANISM: RATTUS RATTUS
US-09-769-097-1

Alignment Scores:

Pred. No.: 1,41e-312 Length: 4369
Score: 3515.00 Matches: 680
Percent Similarity: 73.90% Conservative: 243
Best Local Similarity: 54.44% Mismatches: 266
Query Match: 17.07% Indels: 60
DB: 10 Gaps: 10

US-09-873-409-5 (1-1222) x US-09-769-097-1 (1-4369)

QY 1 MetIleuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
DB 499 ATGCTGCTGGGAATCTGGCGGCATATCCATGGAAATCGCTCCACATATGATGCTG 558
QY 21 ValLeuGlyGluMetSerAspAsnLeuSerGlyCysLeuValGlnThrAsnThrTyr 40
DB 559 GTCTTTGGACATGACAGATAGCTTTGCAAT-----GTAGAAACAACCGTAGTATG 612
QY 41 SerPhePheArgLeuThr----- 46
DB 613 AGTTTCTACAATGCTACAGACATATATGCCAAGCTGGAGGACGAAATGGCCACGTACGCC 672
QY 47 LeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeu 66
DB 673 TACTATTACACGGGCAATGGTGGCGGTGCTCATCGCTTCATCCAGGTTTCACTT 732
QY 67 TrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGlnPhePheHisSerVal 86
DB 733 TGGTGCCTGGCAGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 792
QY 87 LeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGluLeuAsnThrArgMet 106
DB 793 ATGAATCAGGAGATAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 852
QY 107 Thr---AspIleAspIlySerAspGlyIleGlyAspIlySerIleAlaLeuLeuPheGln 125
DB 853 ACAGATGACGTCTCAAAATTAATGAAGGAATGGTGACAAAATTTGGAATGTTCTTTTCAG 912
QY 126 AsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLysGlyTyrPheLeuThr 145
DB 913 GCAATGGCAACATTTTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 972
QY 146 LeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAlaCysSerArgMet 165

DB 973 CTGTGATTTGGCCATCAGCCCTGTCTTGGACTGTGAGCTGGTATTTGGGCAAGATA 1032
QY 166 ValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGlu 185
DB 1033 TTGTCTTCAATTTACTGATGAAGAACTCCAGGCTTATGCAAAAGCTGAGCAGTTGCTGAA 1092
QY 186 GluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGln 205
DB 1093 GAAGTCTTAGCAGCCATCAGAACTGTGATTGCTTTGGAGGACAAAGAAGGAACATTGAA 1152
QY 206 ArgSerPheLeuLeuAsnIleThrArgTyrAlaTyrPheTyrPheProGlnTrpLeuLeu 225
DB 1152 ----- 1152
QY 226 SerCysValLeu***PheValArgTyrThrGlnAsnLeuLeuAspAlaLysAspPheGly 245
DB 1153 -----AGGTACAATAACAAATTTGGAAGAAGCTAAAGAGCTTTGGG 1191
QY 246 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGly 265
DB 1192 ATAAAGAAAGCTATCAGCGCAACATTTCCATGGGTGCAAGCTTTTCTGCTTATCTATGCA 1251
QY 266 ThrTyrGlyLeuAlaPheTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 285
DB 1252 TCATATGCTCTGGCAATCTGGTATGGGACTTCCTCGTCAATCTCAAAAGAA-----TAC 1305
QY 286 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 305
DB 1306 ACTATTGGACAAGTGTCTCACTGCTTTTTTTCTGTAATTAATGGAGCATTCACTGTTGGG 1365
QY 306 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 325
DB 1366 CAGGCATCTCCAAATATTGAAGCCTTCGCCAATGCTAGAGGAGCAGCTTAAGTCTTC 1425
QY 326 GlnValIleAspLysLysProSerIleAspAsnPheSerThrLaglyTyrLysProGlu 345
DB 1426 AGTATAATTGATAAATAAGCCAGTATAGACAGCTTCTCAAAAGAGTGGGCACAAACCCGAC 1485
QY 346 SerIleGlyGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 365
DB 1486 AACATCAAGAAATTTGGAATTCAAATAATATCACTTCAGTTACCCGCTCGAAAGAC 1545
QY 366 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 385
DB 1546 GTTCAGATCTTGAAGGCCTCAACCTGAAGGTGAAGAGCGGCGAGCGTAGCCCTGGTT 1605
QY 386 GlyLeuAsnGlySerGlyLysSerThrValGlnIleLeuGlnArgLeuTyrAspPro 405
DB 1606 GGCAACAGTGGCTGTGGGAAAAGCACAACTGTCCAGCTGTGCAGAGGCTCTACGACCCC 1665
QY 406 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 425
DB 1666 ATAGAGGGCGAGTGTAGTATCGACCGGACAGGACCATCAGGACCATCATGTGAGGTATCTG 1725
QY 426 ArgAspHisIleGlyValValSerGlnProValLeuPheGlyThrThrIleSerAsn 445
DB 1726 CGGGAATCATTTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1785
QY 446 AsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGlu 465
DB 1786 AACATTCGCTATGGCGAGAAAACGTCACCATGGATGAGTAGAAGAGCTGTCAAGGAA 1845
QY 466 AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 485
DB 1846 GCCAATGCCCTATCATTTTCATGAAATTCGCCCAAAATTTGACACCCCTGGTGGTGAG 1905
QY 486 LysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal 505
DB 1906 AGAGGGGGCGAGTGTAGTGGGGGACAGAAACAGAGGATGCCATTCGCCGGCCCTGTC 1965
QY 506 ArgAsnProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 525

Db 1966 CGAACCCCAAGATCCCTTTTGGTGGATGAGCCACAGTCAGCCTTGGACACAGAAAGCGAA 2025
Qy 526 SerAlaValGlnAlaLeuGluValAlaSerLysGlyArgThrThrIleValValAla 545
Db 2026 GCCGTGGTTCAGGCGCTCTGGATAAGGCTAGAGAAGCGGACCACTTGTGATAGCT 2085
Qy 546 HisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeu 565
Db 2086 CACGCTTGCTACAGTTCGCAATGCTGACGTCATTGCTGGTTTGGTGGTGCATT 2145
Qy 566 AlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuVal 585
Db 2146 GTGAGCAGGAATAATCATGATGAGCTCATGAGACGGAAGAAATTAATCTCAAACTGTGC 2205
Qy 586 MetSerGln-----AspIleLysLysAlaAspGluGlnMetGluSer----- 599
Db 2206 ATGACTCAGACAGCAGGAATAATGAATTAAGAAATGAAGCTTGTGAATCTTAAGAC 2265
Qy 600 -----MetThrTyrSerThrGluArgLysThrAsnSerLeu-----ProLeu 613
Db 2266 GGAATTGATAATGGGACATGCTCTTCAAAAGATTTCGGGATCCAGTCTTAATAAGAAGA 2325
Qy 614 HisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThr---Gln 632
Db 2326 TCAACTCGCAAAAGCATCCGTCGGCCACATGATCAAGACGCGGAACCTTAGCACCAGAG 2385
Qy 633 SerLysGluIleSerLeuProGluValSerLeuLysIleLeuLysLeuAsnLysPro 652
Db 2386 GCTCTGGATGACGACGTACCTCCAGCTTCTTTGGCGGATCTTGAAGTTGAATCAACT 2445
Qy 653 GluTrpProPheValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPro 672
Db 2446 GAATGGCTTATTTTGGTGGTGTATTTGTGCCATAATAATGAGGCTTGCAACCA 2505
Qy 673 ValPheSerIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThr 692
Db 2506 GCATCTCCATAATATTTCAAGAGTTGTAGGGGTTTTTACAAAAAATGACACCCCTGAA 2565
Qy 693 Leu---LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCys 711
Db 2566 ATCCAGCGGCAAGACAGCACTGTTTCTTATTTGTTCTGTGATCTTGGGATCATCTCT 2625
Qy 712 PheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMet 731
Db 2626 TTCATTAGCTTTTCCCTCAGGCTTCACATTTGCCAAGCTGGAGAGATCCTCACCAC 2685
Qy 732 ArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGlu 751
Db 2686 CGACTCGGATACATGCTCTTCAAAATCCATGCTGACAGCAGGACATAAGCTGTTGATGAC 2745
Qy 752 LysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGln 771
Db 2746 CCTAAAAACACCAAGAGCGCTGACCACCGGCTTCCCAATGACGCTGCTCAAGTGAAA 2805
Qy 772 GlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSer 791
Db 2806 GGGGCTACAGGCTCTAGGCTGCTGTATTATACCCAGAACATAGCAATCTTGGACAGGC 2865
Qy 792 ValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaPro 811
Db 2866 ATCATCATATCCCTGATCTACGGCTGGCAATTGACATTTTACTCTTAGCAATTTGTTCCC 2925
Qy 812 ValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAsp 831
Db 2926 ATCATGCTATAGCAGGAGTGGTTGAATTAAGAAATGTTGTCTGGACCAAGCGCTCAAGAT 2985
Qy 832 LysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThr 851
Db 2986 AAGAGGNACTAGAGGTTCTGGGAAGATCGCTACAGAGCAATGAATACTTTCGACT 3045
Qy 852 IleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThr 871
Db 3046 GTCTGCTCTTTGACTCGGGAGCAGAAAGTTTGAACATATGATGATGCCAGAGCTTGACATA 3105

Qy 872 GlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHis 891
Db 3106 CCATACAGAAATGCTTTGAAGAAAGCGACGCTTTGGGATCATCTTCTCTTACCAC 3165
Qy 892 AlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAla 911
Db 3166 GCCATGATGATTTCTCTCTATGCTGCTGTTCCGGTTTGTATGCTTACTTGGTGACCA 3225
Qy 912 GlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMet 931
Db 3226 GAATCATGACATTTGAAATGTTCTGTAGTATTCTCAGCTATTGTCTTTGTGGCATG 3285
Qy 932 AlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAla 951
Db 3286 GCAGTGGGCGAGTCAGTTCAATTCGCTCTCTGACTACGCAAGCAAGTCTCGGCATCC 3345
Qy 952 HisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLys 971
Db 3346 CACATCATCAGGATCATTTGAGAAAATCCCTGAGATTCAGAGCTACAGCACGAGGCTTG 3405
Qy 972 LysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCys 991
Db 3406 AAGCCTAATATGTTGAAGGAAATGTAATTTAATGGAGTCATGTTCAACTATCCACC 3465
Qy 992 ArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrVal 1011
Db 3466 CGACCCCAACATCCAGTCTTTCAGGGCTGAGCTAGAGGTGAAGAAAGGCGACGCTG 3525
Qy 1012 AlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeu 1031
Db 3526 GCCTCTGTCGGCAGCAGTGGCTCGGGAAGAGTACAGTGGTCCAGCTGCTTGAGCGCTC 3585
Qy 1032 TyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnVal 1051
Db 3586 TATGACCCCATGTCGCGAAGCAGTGTTCCTAGATGGCAAAATAAGCAATCAATGTC 3645
Qy 1052 GlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSer 1071
Db 3646 CAGTGGCTCGCGCCACCTCGGCAATGTGTCCAGAGGCCCATCTGTTGACTGCAGC 3705
Qy 1072 IleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLys 1091
Db 3706 ATCGCCGAGAAATTCCTACGAGACCAACAGCCGTGTCTGTCTCATAGGAGATCGTG 3765
Qy 1092 GluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsn 1111
Db 3766 AAGGACGCCAAGAGGCGCAACATCCACAGTTCATCGACTCCTGCTGAGAATAACAC 3825
Qy 1112 ThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIle 1131
Db 3826 ACCAGTGGGACAAAGGACTCAGCTGTGGCGGGCAGAGCAGCGCATCGGCATC 3885
Qy 1132 AlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeu 1151
Db 3886 GCGCGCGCCCTCGACAGACGCTCATCTTACTTCTGGATGAAGCAGCATCAGCTCTG 3945
Qy 1152 AspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThr 1171
Db 3946 GATACGGAGAGTGAAGGTCGTCAGGAGCGCTGGCAAAAGCCAGGGAAGCGCGCAC 4005
Qy 1172 CysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeu 1191
Db 4006 TGCATTGTGATCGGCACCGCTGTCCACCATCCAGACCGCAGACTTGTATCGTGTAT 4065
Qy 1192 HisAsnGlyLysLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIle 1211
Db 4066 CAGAACGGCCAGGTCAAGAGACGCGCACCCACAGCAGCTGCTGGGCCAGAAAGGCATC 4125
Qy 1212 TyrPheLysLeuValAsnAlaGlnSer 1220
Db 4126 TATTTCTGATGGTCACTGTGCGAGGCT 4152

RESULT 12

US-10-072-621-2
; Sequence 2, Application US/10072621
; Patent No. US20020169137A1
; GENERAL INFORMATION:

; APPLICANT: Reiner, Peter B.

; APPLICANT: Connop, Bruce P.

; APPLICANT: Pollard, Michelle

; TITLE OF INVENTION: REGULATION OF AMYLOID PRECURSOR PROTEIN EXPRESSION

; TITLE OF INVENTION: BY MODIFICATION OF ABC TRANSPORTER EXPRESSION OR ACTIVITY

; FILE REFERENCE: 100103.402

; CURRENT APPLICATION NUMBER: US/10/072,621

; CURRENT FILING DATE: 2002-02-08

; NUMBER OF SEQ ID NOS: 10

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 2

; LENGTH: 4643

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-072-621-2

Alignment Scores:

Pred. No.:	1,916-312	Length:	4643
Score:	3514.00	Matches:	681
Percent Similarity:	73.92%	Conservative:	246
Best Local Similarity:	54.31%	Mismatches:	263
Query Match:	57.05%	Indels:	64
DB:	9	Gaps:	11

US-09-873-409-5 (1-1222) x US-10-072-621-2 (1-4643)

QY	1	MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu	20
DB	575	ATGGTGGTGGGAACCTTTGGCTGCATCCATCCATGGGCTGGATCTCTCTCATGATGCTG	634
QY	21	ValLeuGlyGluMetSerAsp	31
DB	635	GTGTTTGGAGAAATGACAGATATCTTTGCAATGTCAGGAAATTTAGAAATCTGATGTCA	694
QY	32	GlyCysLeuValGlnThr	43
DB	695	AACATCACTAATAGATGATATATACAGAGGTCTTCATGATATCTGGAGGAGAC	754
QY	44	-----ArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuPheGlyTyr	61
DB	755	ATGACACAGGTATCCCTATATATACAGTGGAAATGGTCTGGGTGCTGGTGTCTTAC	814
QY	62	IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln	81
DB	815	ATTGAGTTTCATTTTGGTCCCTGGCAGCTGGAAGACAAATACACAAATAGAAAACAG	874
QY	82	PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu	101
DB	875	TTTTTTTCATGCTATAATGCGACAGGAGATAGGTGGTTTATGATGTCACGATTTGGGGAG	934
QY	102	LeuAsnThrArgMetThr	120
DB	935	CTTAAACCCGACTTACAGATGATGCTCCAGATTAATGAGGAATTTGGTGCACAAAT	994
QY	121	AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys	140
DB	995	GGAATGTTCTTTCAGTCAATGGCAACATTTTTCATCGGTTTATAGTAGGATTTACACGT	1054
QY	141	GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla	160
DB	1055	GGTTGGGAAGTAACCTTGTGATTTTGGCCATCAGTCTCTGTTTGGACATGTCAGCTGCT	1114
QY	161	AlaCysSerArgMetValIleSerLeuThrSerLysGlnLeuSerAlaTyrSerLysAla	180
DB	1115	GTCGGGCAAGATACATCTTCTTATCTGATTAAGAACTCTTAGCGATGCAAGACT	1174
QY	181	GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln	200

DB	1175	GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGTCATTTGGAGACAA	1234
QY	201	GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe	220
DB	1235	AGAAAAGAACTTGA	1249
QY	221	ProGlnTrpLeuLeuSerCysValLeu	240
DB	1250	-----AGGTACACAAATAATTAGAA	1273
QY	241	AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr	260
DB	1274	GCTAAAGAAATTTGGATGAAGAACTATACAGCAATATTTCTATAGTGTCTGCTTTC	1333
QY	261	PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuLeuLeuAsn	280
DB	1334	CTGCTGATCTATGATCTTATGCTCTGGCTTCTGGTATGGACCACTTGTCTCTCTCA	1393
QY	281	GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer	300
DB	1394	GGGAA-----TATTTCTATGGCAAGTACTCCTGTA	1444
QY	301	SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla	320
DB	1445	GCTTTTATGTTGGACAGCATCTCCAGCATTTGAAATTTGCAATTCAGAGAGGCA	1504
QY	321	AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla	340
DB	1505	GCTTATGAAATCTTCAAGATAATTGATAAAGCCAAAGTATTGACAGCTATTTCGAAGCT	1564
QY	341	GlyTyrLysProGluSerIleGlyThrValGluPheLysAsnValSerPheAsnTyr	360
DB	1565	GGGCAAAACACAGATAATTAAGGAAATTTGGAATTCAGAAATGTTCACTTTCAGTTAC	1624
QY	361	ProSerArgProSerIleLysIleLysGlyLeuAsnLeuArgIleLysSerGlyGlu	380
DB	1625	CCATCTCGAAAAAGATTTAAGATCTTGAAGGGTCTGAACTGAAGGTGACAGTGGGCG	1684
QY	381	ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln	400
DB	1685	ACGTGGCCCTGGTTGGAACAGTGGCTGGGAAGAGACACACAGTCCAGCTGATGCG	1744
QY	401	ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu	420
DB	1745	AGCTCTATGACCCACAGAGGGGATGTCAGTGTGTGATGGACAGGATATTAGGACATA	1804
QY	421	AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly	440
DB	1805	ATGTAAGGTTCCTACGGGAATCATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	1864
QY	441	ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluMetGlu	460
DB	1865	ACCACGATAGTGAACATTCGCTATGCGGTGAAATGTCCACATGGATGAGATTGAG	1924
QY	461	ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn	480
DB	1925	AAAGCTGTCAAGGAAGCAATGCTATGCTTATCATGAACTGCTCATATAATTGAC	1984
QY	481	ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle	500
DB	1985	ACCTGTTGGAGAGAGGGGCCAGTGTGAGTGGTGGGAGAGAGAGAGATCGGCATT	2044
QY	501	AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu	520
DB	2045	GCACGTGCTGCTGTCGCAACCCCAAGATCTCTCTCTGATGAGGCCAGCTGAGCTTG	2104
QY	521	AspSerGluSerLysSerAlaValGlnAlaLeuGluLysAlaSerLysGlyArgThr	540
DB	2105	GACACAGAAAGCAAGCAGTGGTTCAGTGGCTCTGGATAAGGCCAGAAAGGTCCGACC	2164
QY	541	ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu	560
DB	2165	ACCATTTGTATAGCTCATCGTTTGTCTACAGTTTCTGTAATGCTGACGTATCGTGTTC	2224

; CURRENT APPLICATION NUMBER: US/09/866,866A
 ; CURRENT FILING DATE: 2001-08-30
 ; PRIOR APPLICATION NUMBER: 09/584,586
 ; PRIOR FILING DATE: 2000-05-31
 ; PRIOR APPLICATION NUMBER: PCT/US99/11825
 ; PRIOR FILING DATE: 1998-05-27
 ; PRIOR APPLICATION NUMBER: 60/086,988
 ; PRIOR FILING DATE: 1998-05-28
 ; NUMBER OF SEQ ID NOS: 27
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 5
 ; LENGTH: 4189
 ; TYPE: DNA
 ; ORGANISM: Mus musculus
 ; US-09-866-866A-5

Alignment Scores:
 Pred. No.: 1,68e-311 Length: 4189
 Score: 3503.00 Matches: 678
 Percent Similarity: 73.13% Conservative: 242
 Best Local Similarity: 53.90% Mismatches: 262
 Query Match: 56.88% Indels: 76
 DB: 10 Gaps: 11

US-09-873-409-5 (1-1222) x US-09-866-866A-5 (1-4189)

Qy	1	MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu	20
Db	148	ATGATTCTGGGAACCTCGCTGCTATATCCATGGACATTACTTCCCTCTTGATGCTG	207
Qy	21	ValLeuGlyGluMetSerAspAsnLeuLeu	30
Db	208	GTGTTTGGAAACATGACAGATAGTTTACAAAAGCAGAAGCCAGTAGTATTCTGCCAAGCAATT	267
Qy	31	-----SerGlyCysLeuValGlnThrAsnThr-----	39
Db	268	ACTAATCAAAAGTGGACCCACAGTACTGATCATCAGCAACAGCAGTCTGGAGGAAGAG	327
Qy	40	-----TyrSerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAla	56
Db	328	ATGGCCATATACGCCTAC-----TATTACACCGGGATTGGTGTGCTGTG	372
Qy	57	LeuIlePheGlyTyrIleGlnIleSerLeuTyrIleIleThrAlaAlaArgGlnThrLys	76
Db	373	CTCATAGTTGCCATCATCCAGGTTTCACTTGGTGCTGGCAGCTGGACAGACAGATACAC	432
Qy	77	ArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSer	96
Db	433	AAGATTAGGCAGAAAGTTTTCATGTCTATAAATCAGGAGATGGCTGGTTTGATGTG	492
Qy	97	CysAspIleGlyGluLeuAsnThrArgMetThr--AspIleAspLysIleSerAspGly	115
Db	493	CATGATGTTGGGAGCTCAACACCGGCTCACAGATGATGCTCCAAATTAATGACGGA	552
Qy	116	IleGlyAspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAla	135
Db	553	ATTGGTGACAAAATTGGGATGTTTTTTCAGTCCATAAACCAATTTTAGCCGCTTTATC	612
Qy	136	ValGlyLeuValLysGlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIle	155
Db	613	ATAGGATTTTATAAGTGGTTGGAAAGTCAACCCCTGTCAATTTTGGCTGTCAGCCCTCTTATT	672
Qy	156	MetAlaSerAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSer	175
Db	673	GGATTGCTAFTCTCTTGTGGGCAAGGATGATTGACTTCTATTACTAATAAGGAACCTCCAG	732
Qy	176	AlaTyrSerLysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIle	195
Db	733	GCTTATGCAAAACCTGGAGCAGTTGCTGAAGAAGTCTTAGCAGCCATCAGAACTGTGATT	792
Qy	196	AlaPheArgAlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyr	215
Db	793	GCCTTTGGAGGACACAGAAGGAACCTTGA-----	822

Qy	216	AlaTyrPheTyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThr	235
Db	823	-----AGGTACAAT	831
Qy	236	GlnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSer	255
Db	832	AAAAATTTAGAAAGAGCTAAAAATGTTGCATAAAGAAAGATATCACAGCCGACATTCG	891
Qy	256	LeuGlyAlaValTyrPheMetAsnGlyThrTyrGlyLeuAlaPheTyrPyrGlyThr	275
Db	892	ATAGGATTCCTCTCAAAATGAA-----TATTCTATTGGAGAAGTGTCTACTGTCTCTTC	951
Qy	276	SerLeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePhe	295
Db	952	TCCTTGGTCTCTCAAAATGAA-----TATTCTATTGGAGAAGTGTCTACTGTCTCTTC	1005
Qy	296	SerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAla	315
Db	1006	TCTATTTTGTGGGACTTTTATGATTGGACACTTGGCCCCCAACATAGAAGCTTTTGA	1065
Qy	316	IleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAsp	335
Db	1066	AACGACAGGGGGCAGCCTTTCAAAATCTTCAAGATAATTGATAACGAGCCAGCATTTGAC	1125
Qy	336	AsnPheSerThrAlaGlyTyrLysProGluSerIleGlyThrValGluPheLysAsn	355
Db	1126	AGCTTCTCAACAAAGGGCTACAAACCCAGACAGTATTAATGGGAACTTAGAGTTTAAAT	1185
Qy	356	ValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArg	375
Db	1186	GTTCACTTCAATACCATCGAGAACGAAAGTTCAGATCTTGAAGGCGCTCAATCTGAAG	1245
Qy	376	IleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrVal	395
Db	1246	GTGAGAGCGGACAGACGGTGGCTTGGTTGGCAACAGTGGCTGTGGAAGAACCACT	1305
Qy	396	ValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsn	415
Db	1306	GTCCAGCTCATGCAGAGGCTCTACGACCCCTGGAGGGCGTGTGTCAGTATCGAGACAA	1365
Qy	416	AspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGlu	435
Db	1366	GACATCAGAACCAATCAATGTAGGTATCTGAGGGAGATCATTTGGTGTGTGAGTACGAA	1425
Qy	436	ProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspValThr	455
Db	1426	CCTGTCTGTTTCCACCCAGCATGCCGAGAACATTCGCTATGGCCGAGAAGATGTACC	1485
Qy	456	AspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPhe	475
Db	1486	ATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAATGCCTATGACTTTCATGAAGCTG	1545
Qy	476	ProAsnLysPheAsnThrLeuValGlyLysGlyAlaGlnMetSerGlyGlnLys	495
Db	1546	CCCCACCAATTTACACCCCTGGTTGGTGAGAGAGGGCGCAGCTGAGTGGGGACAGAA	1605
Qy	496	GlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGlu	515
Db	1606	CAGAGAAATCGCATTCGCCGGGCGCTGGTCCGCAATCCCAAGATCCTTTTGTGGACGAG	1665
Qy	516	AlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAla	535
Db	1666	GCCACTTCAGCCCTGGATACAGAAAGTGAAGTGTGTGTGTCAGGCGCCACTGGATAAGCT	1725
Qy	536	SerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAsp	555
Db	1726	AGAGAGGCGGACCAACCATTTGTGATAGTCTATCGCTTGTCTACAGTCTGTATGTCTGAC	1785
Qy	556	LeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMet	575
Db	1786	GTCATTGCTGGTGTGTGATGGTGTGTGTCATTGTGGACGAAGGAATCATGATGAGCTCATG	1845

; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: 60/086,988
; PRIOR FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 4788
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-866-866A-7

Alignment Scores:

Pred. No.: 4,776-311 Length: 4788
Score: 3499.00 Matches: 668
Percent Similarity: 73.13% Conservative: 252
Best Local Similarity: 53.10% Mismatches: 264
Query Match: 56.81% Indels: 74
DB: 10 Gaps: 9

US-09-873-409-5 (1-1222) x US-09-866-866A-7 (1-4788)

```
Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 148 ATGCTGGTGGAACTCTGGCTGCTATATCCATGGAGTGGCGCTCCACCTATATGATGCTG 207
Qy 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsn----- 38
Db 208 ATCTTTGGTGACATGACAGATAGCTTTGGCAAGTGTAGGAAACGCTCTCTAAACACAGTACT 267
Qy 39 -----Thr 39
Db 268 AATATGATGAGCGCCGATAAAGAGCGCATGTTTGGCCAACTGGGAGGAAGAAATGACCAACG 327
Qy 40 TyrSerPhePheArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePhe 59
Db 328 TAGCCCTAC-----TATTACCCGGATTGGTCTGGTGTGCTCATAGTT 372
Qy 60 GlyTyrIleGlnIleSerLeuThrIleIleThrAlaAlaArgGlnThrLysArgIleArg 79
Db 373 GCCTACATCCAGGTTTCATTTTGGTGGCGCTGAGAGTGTTCCTCAAAATTAATGAAGGAATTTGGTGAC 552
Qy 80 LysGlnPhePheHisSerValLeuAlaGlnApeIleGlyTrpPheAspSerCysAspIle 99
Db 433 CAGAAGTTTTTTTCATGCTTAATGAATGAATGAGAGATAGGCTGGTTGATGTGCATGACGTT 492
Qy 100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
Db 493 GGGGAGCTCAACCCCGCTCAGATGATGTTTCCAAATTAATGAAGGAATTTGGTGAC 552
Qy 119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
Db 553 AAAATCGGAATGTTCTTCAGGCAATGGCAACATTTTGGTGTATTATATAGATTT 612
Qy 139 ValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
Db 613 ACCCGTGGGTGGAAGCTAAACCTTGTGATTTTGGCCATCAGCCCTGTCTTGGCATGTCA 672
Qy 159 AlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
Db 673 GCTGGTATTGGCAAGATATTGCTTCATTACTGATAAGGAACCTCCATGCTTATGCA 732
Qy 179 LysAlaGlyValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198
Db 733 AAAGCTGGAGCAGTCTGTAAGAGTCTTAGCAGCCATCAGAACTGTGATGGCTTGA 792
Qy 199 AlaGlnGlyGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPhe 218
Db 793 GGACAAAGAGGAAGTCTGAA-----813
Qy 219 TyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeu 238
Db 814 -----AGGTACAATAACAACCTTG 831
```

```
Qy 239 LysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAla 258
Db 832 GAAGAAAGCTAAAAGGCTGGGATAAAGAAAGCTATCACGCCAACATCTCCATGGGTGCA 891
Qy 259 ValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIle 278
Db 892 GCTTTTCTCTTATCATCATCATATGCTCTGGCAATCTTGGTATGGGACTTCTCTTGGTC 951
Qy 279 LeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIle 298
Db 952 ATCTCCAAAGAA-----TACTCTATTGGCAAGTCTCCTCTCTCTTTCCTGTTA 1005
Qy 299 HisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArg 318
Db 1006 ATTGGAGCATTCAGTGTGGCAGGCAATCTCCAAATATTGAAGCCTTCGCCAATGCACGA 1065
Qy 319 GlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSer 338
Db 1066 GGAGCAGCTTATGAAGTCTTCAAAATAATTGATAATAAGCCAGTATAGACAGCTTCTCA 1125
Qy 339 ThrAlaGlyTyrLysProGluSerIleGlyThrValGluPheLysAsnValSerPhe 358
Db 1126 AAGAGTGGGCACAAACACACACATACAAAGNAATCTCGAATTTAAGAAATATTCACTTC 1185
Qy 359 AsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSer 378
Db 1186 AGTTACCCATCTCGAAAAGAAAGTTTCAAGTCTTGAAGGGCTCAATCTGAAGGTGAAGAGC 1245
Qy 379 GlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeu 398
Db 1246 GGACAGACGCTGGCTGGTGGCAACAGTGGCTGTGGAAAGACACAACTGTCTCAGCTG 1305
Qy 399 LeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArg 418
Db 1306 ATGCAAGGCTCTACAGCCCTTAGATGCGATGCTGATGCGACAGGACATCAGA 1365
Qy 419 AlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeu 438
Db 1366 ACCATCAATGTGAGGTATCTGAGGGAGATCATTTGGTGTGTGAGTCAGGAACCTGTGCTG 1425
Qy 439 PheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGlu 458
Db 1426 TTTGCCACCAACCATGCCGAGAACATTCCTATGCGCGAGAAAGATGTCACCATGGATGAG 1485
Qy 459 MetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLys 478
Db 1486 ATGAGAAAGCTGTCAAGGAACCAATGCTATGATCTCATGAAGACTGCCCCACCAA 1545
Qy 479 PheAsnThrLeuValGlyLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIle 498
Db 1546 TTTGACACCCCTGGTTGGTGAGAGAGGGGCGCAGTGGAGTGGGGACAGAAACAGAAATC 1605
Qy 499 AlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSer 518
Db 1606 GCATTTGCCCGGCGCTGTCGCAATCCCAAGATCTTTTGTGGACGAGGCCACTCA 1665
Qy 519 AlaLeuAspSerGluSerLysSerAlaValGlnAlaLeuGluLysAlaSerLysGly 538
Db 1666 GCCCTGATACAAAGTGAAGCTGTGGTTGAGGCGCGACTGGATGAAGCTAGAGAGGC 1725
Qy 539 ArgThrThrIleValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleVal 558
Db 1726 CGGACCAACATTTGTATAGTCTATCGTGTCTACCGTTCTGTAATGCTACGCTCATTTGCT 1785
Qy 559 ThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArg 578
Db 1786 GGTTTTGTATGGTGTCTATGTTGGCAAGGAATCATGATGAGCTCATGAGAGAAAG 1845
Qy 579 GlyLeuTyrTyrSerLeuValMetSerGln-----588
Db 1846 GGATTTACTTCAAACTTGTGTCATGACACAGCAGCAGGAAATGAAATTAAGTAAAT 1905
Qy 589 AspIleLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThr 608
```



```
Db 1906 GAAGCTTGTAAATCTAAGGATGAATTAATTTAGACATGTCCTTCAAAAGATTGAGGA 1965
Qy 609 AnSerserProLeuHisSerValLysSerIleLysSerAspPheIle 624
Db 1966 TCCAGTCTAATAAGAAGA-----AGATCAACTCGCAAAAGCATCTGTGGACACACATGAC 2019
Qy 625 ---AspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeu 643
Db 2020 CAAGACAGGAAGCTTAGTACCAAGAGGCCCTCGATGATGATGATGATGATGATGATGAT 2079
Qy 644 LeuLysIleLeuLysLeuAsnLysProGluTrpPheValValLeuGluThrLeuAla 663
Db 2080 TGCCGATCTGAAGTGAATCAACTGAATGCGCTTATTTTGTGGTGGTATATTCTGT 2139
Qy 664 SerValLeuAsnGlyThrValHisProValPheSerIlePheAlaLysIleLeuThr 683
Db 2140 GCCATAATAATGGAGCTTACAGCCAGCAATCTCCGTAATATTTTCAAAAGTTGTAGG 2199
Qy 684 MetPheGlyAsn---AnAspLysThrThrLeuLysHisAspAlaGluIleLysSerMet 702
Db 2200 GTTTTACAAATGGTGGCCCTGAAACCCAGCGGAGAACAGCAACTTGTGTTTCTCTG 2259
Qy 703 IlePheValIleLeuGluValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyr 722
Db 2260 TTGTTTCTGATCTCTGGGATCATTTCTTCAATACATTTTCTTCTCAGGGCTTCACATTT 2319
Qy 723 GlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeu 742
Db 2320 GGCAAACTGGAGATGCTTCAAGTCAAGGGGGTACAGGGTCTAGGGTGTGTGATTTTC 2379
Qy 743 TyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrIle 762
Db 2380 AGACAGGATGTAGCTGTTTGTATGACCTTAAACACACCCAGGAGCACTGACCAACAGG 2439
Qy 763 LeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThr 782
Db 2440 CTGCCAACATGCTGCTCAAGTCAAGGGGGTACAGGGTCTAGGGTGTGTGATTTTC 2499
Qy 783 GlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyrGluMet 802
Db 2500 CAGAACATCAAACTTTGGGACAGGATCATCATATCCCTATCTATGCTGGCAACTA 2559
Qy 803 ThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAla 822
Db 2560 ACATTTTACTTCTAGCAATTTGACCCATCATCTGCGATAGCTGGAGTGGTGAATGAAA 2619
Qy 823 AlaMetThrGlyPheAlaAsnLysAspLysGlnLeuLysHisAlaGlyLysIleAla 842
Db 2620 ATGTTGTCTGGCAAGCACTGAAAGATAAGAGGAACTAGAAAGTTCCTGGAAGATTGCT 2679
Qy 843 ThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGlu 862
Db 2680 ACGGAACAATTAAGAACTTCCGACATGTTGCTCTTTTGTACTCGGGAGCAAGTTTGAA 2739
Qy 863 GlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIle 882
Db 2740 ACCATGTATCCAGAGCTTGCAGATACCATACAGAAATCCGATGAAGAAGCACAGGTG 2799
Qy 883 IleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPhe 902
Db 2800 TTTGGGATCAGTTCTCTTCCACCGGCCCATGATGATGATGATGATGATGATGATGATGAT 2859
Qy 903 ArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleVal 922
Db 2860 CGGTTCGGTCCCTACTTGGTGACACAACTCATGATCTTTGAAATGTTCTGTGTAGTA 2919
Qy 923 PheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGlu 942
Db 2920 TTCTCAGCTATGCTTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2979
Qy 943 TyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuGluLysLysProAsn 962
```

```
Db 2980 TATGCGAAAGCAACAGTGTGATGATCCCATCATCATCATCATCATCATCATCATCATCAT 3039
Qy 963 IleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPhe 982
Db 3040 ATTGACAGCTACAGCAGCAGCAGCCTTAAAGCCGAATATGTTGGAAGAAATGTGCAATTT 3099
Qy 983 ArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSer 1002
Db 3100 AGTGGAGTGTGTTCAACTATATCCACCCAGCAGCATCCAGTGTCTTCCAGGGCTGAGC 3159
Qy 1003 LeuSerIleGluArgGlyLysThrValAlaPheValGlySerGlyCysGlyLysSer 1022
Db 3160 CTTGAGTGAAGAGGCCAGAGCTGTGGCCCTGTGGGAGCAGTGTGCTGCGGAAGAGC 3219
Qy 1023 ThrSerValGlnLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAsp 1042
Db 3220 ACAGTGTCTCAGCTGCTCGAGCGCTTCTACGACCCCATCGCTGGATCAGTGTCTTAGAT 3279
Qy 1043 GlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValPro 1062
Db 3280 GGCAGAAATAAAGCACTGAATGTCCAGTGGCTCCGAGCACAGCTGGGCATTTGTGTC 3339
Qy 1063 GlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSer 1082
Db 3340 CAAGACCCCATCTCTTCTGACTGCGAGCATCGCAGAGAACTTCCCTACGAGCAACAGC 3399
Qy 1083 ArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPhe 1102
Db 3400 CGGGTGTCTTATGAGGAGATTGTGAGGGCAGCCAGGAGGCCAACATCCACCAAGTTC 3459
Qy 1103 IleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSer 1122
Db 3460 ATGACTGCTCTCTCTATTAATACACACAGTAGTAGGAGCAAAAGGCATCTCAGCTGTGC 3519
Qy 1123 GlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuGlnLysProLysIleLeu 1142
Db 3520 GGTGGGCAAGCAGCGCATCGCCATCGCAGCGCCCTCGTACAGCAGCTCATATTTTA 3579
Qy 1143 LeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAla 1162
Db 3580 CTTCTGAGCAAGCAACATCAGCTCTGGATACAGAAAGTGAAAGGTTGTCCAGGAAGCG 3639
Qy 1163 LeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIle 1182
Db 3640 CTGGAACAAGCCAGGAGCGCCGACCTGCTCATTTGTATCGCTCAGCCCTGTCCACCATC 3699
Qy 1183 GlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHis 1202
Db 3700 CAGACCGGACTTGTATCGTGTGATTTCAGACGGCAAGGTCAAGGAGCAGCGACCCAC 3759
Qy 1203 GlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
Db 3760 CAGCAGCTCTGCGCAGAGGGCATCTACTTCTCATGTTGTCAGTGTGAGGCT 3813

RESULT 15
US-09-880-107-2299
; Sequence 2299, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: Patentin Ver. 2.1
```



```

; SEQ ID NO 2299
;
; LENGTH: 3924
;
; TYPE: DNA
;
; ORGANISM: Homo sapiens
;
; FEATURE:
;
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 M23234
US-09-880-107-2299

```

Alignment Scores:			
Seq. No.:	1-596-309	Length:	3924
Score:	3481.00	Matches:	676
Percent Similarity:	72.82%	Conservative:	235
Best Local Similarity:	54.04%	Mismatches:	276
Query Match:	56.52%	Indels:	64
DB:	10	Gaps:	9
US-09-873-409-5 (1-1222) x US-09-880-107-2299 (1-3924)			
Qy	1	MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu	20
Db	201	NYGTCGTGGTACCATCATCGGCATAGCTCAGGATCAGGTCCTCCCTCATGATGATA	260
Qy	21	ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTy	40
Db	261	GTATTTGGAGAGATGACTGCAAAATTTGTGAT-----ACTGCAGGAAACTTC	308
Qy	41	SerPhe-----PheArgLeuThrLeu-----	47
Db	309	TCCTTTCCAGTGAACCTTTCTTGTCTGCTTAAATCCAGGCAGAAATTCGAAAGAGAA	368
Qy	48	-----TYTYrValGlyIleGlyValAlaAlaLeuIlePheGlyTYr	61
Db	369	ATGACTAGATATGCATATTACTACAGATGGGTGGGATCTTGTGCTGCCTAT	428
Qy	62	IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln	81
Db	429	ATACAAAGTTTCATTTTGGACTTTGGCAGCTGGTCGACAGATCAGGAAAATTAGGCAGAA	488
Qy	82	PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu	101
Db	489	TTTTTTTCATGCTATTCTACGACAGGAAATAGGATGGTTGCATCATCAATGACACCATGAA	548
Qy	102	LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle	120
Db	549	CTCAATACGGCGCTACAGATGACATCTCCAAATCAGTGAAGAAATTTGGTGCACAGTT	608
Qy	121	AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys	140
Db	609	GGAAATGTTCTTTCACAGCAGTAGCCAGTCTTTTTCAGGAGTTCATAGTGGGATTCATCAG	668
Qy	141	GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla	160
Db	669	GGATGGAAGCTCACCTTGTGATAATGGCCCATCAGCCCTATTCTAGGACTCTCTGACGC	728
Qy	161	AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla	180
Db	729	GTTTGGCAAGATACCTCTCGGCATTAGTGACAAAGACTAGCTGCTTATGCAAAAGCA	788
Qy	181	GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln	200
Db	789	GGCGCGCTGGCAGAAGAGGCTCTGGGGGCCATCAGGACTGTGATAGCTTTTCGGGGGCCAG	848
Qy	201	GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyAlaTrpPheTyPhe	220
Db	849	AACAAGAGCTGGAA-----AGGTATCAGAAACATTTAGAAAAT	863
Qy	221	ProGlnTrpLeuLeuSerCysValLeu**PheValArgTyThrGlnAsnLeuLysAsp	240
Db	864	-----AGGTATCAGAAACATTTAGAAAAT	887
Qy	241	AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTy	260
Db	888	GCCAAAGAGATTGGAAATTAAGAAAGCTATTTCAGCAAAACATTTCCATGGGTATTCGCTTC	947

Qy	261	PhePheMetAsnGlyThrTyrglyLeuAlaPheTrpTyrglyThrSerLeuIleLeuAsn	280
Db	948	CTGTTAATATATGTCATATGCACTGGCCCTCTGGTATGATGCATCCACTAGTCATATCA	1007
Qy	281	GlyCUpProGlyTyThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer	300
Db	1008	AAAGAA-----TATACATATTGGAATGCAATGACAGTTTTTTTTTCAATCCTTAATTGG	1061
Qy	301	SerTyrcysIleGlyAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla	320
Db	1062	GCCTTCAGTGTGGCCAGGCTGCCCATGTATTGATGCTTTTGCCAAATGCAAGAGGAC	1121
Qy	321	AlaPheHisIlePheGlnValIleAspIysPysProSerIleAspAsnPheSerThrAla	340
Db	1122	GCATATGTCATCTTTGATATTATTGATAATAATCTTAATAATTGACAGTTTTTTCAGAGAGA	1181
Qy	341	GlyTyIysProGluSerIleGluGlyThrValGluPheIysAsnValSerPheAsnTyr	360
Db	1182	GSACAAACACAGACAGCATCAAGGGAATTTGGAGTTTCAATGATGTTCACCTTTCTTTC	1241
Qy	361	ProSerArgProSerIleIysIleLeuIysGlyLeuAsnLeuAlaIleIysSerGlyGlu	380
Db	1242	CCCTTCGAGCTAACGCTCAAGATCTTTGAAGGCCCTCAACCTGAAGGTGCAGATGGGCAG	1301
Qy	381	ThrValAlaLeuValGlyLeuAsnGlySerGlyIysSerThrValValGlnLeuLeuGln	400
Db	1302	ACGGTGGCCCTGGTTGGAAGTAGTGGCTGTGGGAAGACACAAACGGTCCAGCTGATACAG	1361
Qy	401	ArgLeuTyraAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu	420
Db	1362	AGGCTCTATGACCTGATGAGGCAATTAATCACTGTATGGCAGGATATTAGGAACCTTT	1421
Qy	421	AsnValArgHisTyraArgAspHisIleGlyValValSerGlnGluProValLeuPheGly	440
Db	1422	AATGTAAACTATCTGAGGAAGATCATGTGTGTGGTGTGAGTCAGAGCCGGTCTGTTTTC	1481
Qy	441	ThrIleSerAsnAsnIleIysTyrglyArgAspAspValThrAspGluGluMetGlu	460
Db	1482	ACCACAATTCGTGAAATATTTGTTATGGCGTGGAAATGTAACCATGGATGAGATAAAG	1541
Qy	461	ArgAlaAlaArgGluAlaAsnAlaTyraAspPheIleMetGluPheProAsnIysPheAsn	480
Db	1542	AAAGCTGTCAAAGAGGCCAACGCTATGAGTTATCATGAAATATACCACAGAAATTTGCAC	1601
Qy	481	ThrLeuValGlyGluIysGlyAlaGlnMetSerGlyGlyGlnIysGlnArgIleAlaIle	500
Db	1602	ACCCCTGGTTGGAGAGAGGGGCCACGCTGAGTGTGTGGCAGACAGCAGGATCGCCATT	1661
Qy	501	AlaArgAlaLeuValArgAsnProIysIleLeuIleLeuAspGluAlaThrSerAlaLeu	520
Db	1662	GCACGTGCCCTGTTGCGAACCCCAAGATCTCTTCGTGGATGAGCCACCGTCAGCATTG	1721
Qy	521	AspSerGluSerIysSerAlaValGlnAlaAlaLeuGluIysAlaSerIysGlyArgThr	540
Db	1722	GACACAGAAAGTGAAGCTCAGGTACAGGCAGCTCTGGATTAAGCCACAGAAAGGCGGACC	1781
Qy	541	ThrIleValValAlaHisArgIleuSerThrIleArgSerAlaAspLeuIleValThrLeu	560
Db	1782	ACCATTTGTATGACACACCGACTGTCTACGGTCCCAAAATGCGATGTTCATCGCTGGGT	1841
Qy	561	LysAspGlyMetLeuAlaGluIysGlyAlaHisAlaGluLeuMetAlaIysArgGlyLeu	580
Db	1842	GAGATGGAGTAAATTGTGAGCAGAGGAGCCACAGGAACTGATGTAAGAGGAAGGGGTG	1901
Qy	581	TyrTyrsSerLeuVal-----MetSerGlnAspIleIys	591
Db	1902	TACTTCAAACTTGTCAACATGCAGACATCAGGAAGCCAGATCCAGTCAGAGAAATTTGAA	1961
Qy	592	LysAlaAspGluGlnMetGluSerMetThrTyrsThrGluArgIysThrAsnSerLeu	611
Db	1962	CTAAATGATGAAAGGCTGCCATAGAAATGGCCCAAAATGGCTGGAATCTCGCCTATTT	2021


```
QY 612 ProLeuHisSerVallySerIleLys---SerAspPheIleAspLysAla---GluGlu 629
Db 2022 AGGCATTCTACTCAGAAAACCTTTAAATAATCAAAATGTGTGAGAGAGCGTGTATGTG 2081
QY 630 SerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeu 649
Db 2082 GAAACCGATGGACTTGAAGCAATGTGCCACCAAGTGTCTTCTGAAGGTCCTGAAACTG 2141
QY 650 AsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThr 669
Db 2142 AATAAACAAGATGGCCCTACTTGTGTGGGAACAGTATGTGCCATGGCAATGGGGG 2201
QY 670 ValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAsp 689
Db 2202 CTTGACCGCGCAATTTTCAGTCATATCTCAGAGATCATAGCAGATTTTGGACGAGCGAT 2261
QY 690 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 709
Db 2262 GATGCAAGTGAAGCAGCAGAAAGTGCACATATTTCTTTTGAATTTCTTATTTCTGGGAAT 2321
QY 710 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 729
Db 2322 ATTTCTTTTACTTCTCTCTCAGGGTTTCAAGTGTGGGAAGCTGGCAGATCCCTC 2381
QY 730 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 749
Db 2382 ACCAGAAGACTGGCGTCAATGGGCTTTTAAAGCAATGTGAAGCAGGACATGAGCTGGTT 2441
QY 750 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 769
Db 2442 GATGACCAATAAACAAGTACTGTGTGCCTTTTACAGACTTGCACAGATGCTGCCCAA 2501
QY 770 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 789
Db 2502 GTCCAGGACCCACAGCAACAGGTTGGCTTTAATGCCACAGATATAGTAACCTTGA 2561
QY 790 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle 809
Db 2562 ACTGGTATTATCATATCATTTATCTACGGTTGGCGTAACCTTATTTGCTATTAGCATT 2621
QY 810 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 829
Db 2622 GTTCCAAATTATGTGTGAGAAATTTGTAATGAATGAATTTGTGGTGAATAATGCCAAA 2681
QY 830 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 849
Db 2682 AGAGATAAATAAAGACACTGGAGCTGTGGAAGATGCCACAGAGCAATAGAAATATT 2741
QY 850 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 869
Db 2742 AGGACAGTTGTGCTTTGACCCAGGAAGAAAATTTGAATCAATGTATGTGAAAATTTG 2801
QY 870 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 889
Db 2802 TATGGACCTTACAGGAATTTCTGCAAGAGGCACACATCTATGGAAATCTTTTAGTATC 2861
QY 890 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 909
Db 2862 TCACAGCATTTATGATATTTTCTATGCGGTGTGTTTCGATTTGGTCATATCTCATTT 2921
QY 910 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 929
Db 2922 GTGAATGGACATTCGCTTTCAGAGATGTTATTTCTGTGTGTTTCTGCAATTTGATTGGT 2981
QY 930 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 949
Db 2982 GCAGTGGCTCTAGGACATGCGCAGTTCTATTTGCTCCAGACTATGCTAAGCTAAGCTAGCTCT 3041
QY 950 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 969
Db 3042 GCAGCCCATTTATTCATGCTGTTTGAAGACCAACCTCTGATTGACAGCTACAGTGAAGAG 3101
QY 970 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 989
```

```
Db 3102 GGGCTGAACCCCTGATAAATTTGAAGGAATAATAACATTTAATGAAGTCGTGTCAACTAT 3161
QY 990 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 1009
Db 3162 CCCACCCGAGCAACAGTCGACAGTCTTTCAGGGGCTGAGCTGGAGGTGAAGAAGCCAG 3221
QY 1010 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuGln 1029
Db 3222 ACATAGCCCTGTGGGACAGCAGTGGCTGTGGGAAGACACGCTGTCTCCAGCTCTCTGGAG 3281
QY 1030 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 1049
Db 3282 CGGTCTTACGACCCCTTGGCGGGACAGTGTCTCTGATGGTCAAGAAGCAAGAAACTC 3341
QY 1050 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 1069
Db 3342 ATGTCCAGTGGCTCAGACTCACTCGGAATCGTGTCTCAGGAGCTATCTCTATTGTGAC 3401
QY 1070 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 1089
Db 3402 TGCAGCATTTGCCAGAAATATTCCTATGGAGACAACACAGCCGGGTTGTATCACAGGATGA 3461
QY 1090 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluLeuProGluLys 1109
Db 3462 ATTTGAGTGCAGCCCAAGCTGCCAACATACATCTTTTCATCGAGACGTTACCCCCCAAA 3521
QY 1110 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 1129
Db 3522 TATGAACAAGAGTGGAGATTAAGGGACTCAGCTCTCAGGAGGTCAAAAACAGAGGATT 3581
QY 1130 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 1149
Db 3582 GCTATTGCCCGAGCCCTCATCAGACAACCTCAATCTCTCTGTGGATGAAGCTACATCA 3641
QY 1150 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 1169
Db 3642 GCTCTGGATACTGAAAGTGAAAGGTTGTCCAAGAGCCCTTGACAAAGCCAGAGAGGC 3701
QY 1170 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 1189
Db 3702 CGCACCTGTATGTGATTTGCTCACCGCTGTCCACCATCCAGATGCAGACTTAATAGTG 3761
QY 1190 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 1209
Db 3762 GTGTTTCAGAAATGGGAGAGTCAAGGAGCATGGCAGCATCAGCAGCTGTCTGGCAGAAA 3821
QY 1210 AspIleTyrPheLysLeuValAsnAlaGlnSer 1220
Db 3822 GGCATCTATTTTTCATATGTGTGTCAGTGTCCAGGCT 3854
```

Search completed: March 31, 2003, 16:09:09

Job time : 711.798 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 31, 2003, 02:24:20 ; Search time 178.956 Seconds
(without alignments)
2094.145 Million cell updates/sec

Title: US-09-873-409-5
Perfect score: 6159
Sequence: 1 MILGILASLVNACGLPLML.....QELLENRDYIFKLVAQSQV 1222

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DEV=xlh

-MODE=frame+ p2n.model -DEV=xlh
-Q/cgn2_1/USFTO_spool/US09873409/runat_27032003_115422_19259/app_query.fasta_1.7544
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blom62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09873409 @CGN 1.1 578 @runat_27032003_115422_19259 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:
1: /cgn2_6/prodata/2/ina/5A COMB.seq:*
2: /cgn2_6/prodata/2/ina/5B COMB.seq:*
3: /cgn2_6/prodata/2/ina/6A COMB.seq:*
4: /cgn2_6/prodata/2/ina/6B COMB.seq:*
5: /cgn2_6/prodata/2/ina/PCUS COMB.seq:*
6: /cgn2_6/prodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3521.5	57.2	4646	1 US-08-181-471-2	Sequence 2, Appli
2	3521.5	57.2	6505	1 US-08-793-610-5	Sequence 5, Appli
3	3521.5	57.2	9318	2 US-08-793-610-6	Sequence 6, Appli
4	3520	57.2	4264	2 US-08-784-649A-1	Sequence 1, Appli
5	3520	57.2	4264	2 US-08-784-649A-5	Sequence 5, Appli
6	3519.5	57.1	4669	6 5206352-3	Patent No. 5206352
7	3503.5	56.9	4669	2 US-08-752-447-1	Sequence 1, Appli
8	3503.5	56.9	4669	4 US-09-316-167-1	Sequence 1, Appli
9	3494	56.7	4233	3 US-09-120-513-1	Sequence 1, Appli
10	3494	56.7	4233	4 US-09-450-105-1	Sequence 1, Appli
11	3473.5	56.4	4669	2 US-08-583-276-18	Sequence 18, Appli
12	2280	37.0	2726	1 US-08-461-823-1	Sequence 1, Appli

13	2259	36.7	4002	2 US-08-996-545-1	Sequence 1, Appli
14	2259	36.7	4002	2 US-08-996-545-3	Sequence 3, Appli
15	2259	36.7	4002	4 US-09-328-320-1	Sequence 1, Appli
16	2259	36.7	4002	4 US-09-328-320-3	Sequence 3, Appli
17	2250	36.5	4047	2 US-08-612-734B-1	Sequence 1, Appli
18	2187.5	35.5	4800	2 US-08-612-734B-3	Sequence 3, Appli
19	2143.5	34.8	3924	1 US-08-395-246C-1	Sequence 1, Appli
20	2100.5	34.1	4224	1 US-08-612-521-1	Sequence 1, Appli
21	1722	28.0	6143	1 US-08-612-521-3	Sequence 3, Appli
22	1680.5	27.3	3924	2 US-08-996-644-3	Sequence 3, Appli
23	1680.5	27.3	3924	3 US-09-352-552-3	Sequence 3, Appli
24	1680.5	27.3	3927	2 US-08-996-644-1	Sequence 1, Appli
25	1680.5	27.3	3927	3 US-09-352-552-1	Sequence 1, Appli
26	1646.5	26.7	3909	1 US-08-232-537-1	Sequence 1, Appli
27	1636	26.6	3792	4 US-09-351-224E-10	Sequence 10, Appli
28	1501.5	24.4	3999	4 US-09-351-224E-9	Sequence 9, Appli
29	1106.5	18.0	13188	4 US-08-961-527-70	Sequence 70, Appli
30	1047.5	17.0	7760	4 US-08-961-527-63	Sequence 63, Appli
31	989	16.1	441529	4 US-09-103-840A-1	Sequence 1, Appli
32	987	16.0	4403765	4 US-09-103-840A-2	Sequence 2, Appli
33	940	15.3	6492	4 US-08-961-527-188	Sequence 188, Appli
34	938	15.2	7186	4 US-08-961-527-39	Sequence 39, Appli
35	738	12.0	2376	1 US-08-394-880B-1	Sequence 1, Appli
36	708.5	11.5	1749	4 US-09-134-001C-1893	Sequence 1893, Ap
37	695	11.3	5889	1 US-08-463-092B-5	Sequence 5, Appli
38	695	11.3	5889	2 US-08-462-109A-5	Sequence 5, Appli
39	695	11.3	5889	2 US-08-460-907B-5	Sequence 5, Appli
40	695	11.3	5889	3 US-08-463-179A-5	Sequence 5, Appli
41	695	11.3	5889	3 US-08-461-384B-5	Sequence 5, Appli
42	687	11.2	5011	1 US-08-463-092B-3	Sequence 3, Appli
43	687	11.2	5011	2 US-08-462-109A-3	Sequence 3, Appli
44	687	11.2	5011	2 US-08-460-907B-3	Sequence 3, Appli
45	687	11.2	5011	3 US-08-463-179A-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-08-181-471-2
; Sequence 2, Application US/08181471
; Patent No. 5641508
; GENERAL INFORMATION:
; APPLICANT: Li, Lingna
; APPLICANT: Lishko, Valeryi K.
; TITLE OF INVENTION: METHOD FOR DELIVERING BENEFICIAL
; TITLE OF INVENTION: COMPOSITIONS TO HAIR FOLLICLES
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Thomas Fitting
; STREET: 12526 High Bluff Drive, Suite 300
; CITY: San Diego
; STATE: CA
; COUNTRY: USA
; ZIP: 92130
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/181,471
; FILING DATE: 13-JAN-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/041,553
; FILING DATE: 02-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: ANT0029P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-792-3680
; TELEFAX: 619-792-8477

INFORMATION FOR SEQ ID NO: 2:

* SEQUENCE CHARACTERISTICS:
 LENGTH: 4646 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 425..4267
 US-08-181-471-2

Alignment Scores:

Pred. No.: 0 Length: 4646
 Score: 3521.50 Matches: 681
 Percent Similarity: 73.92% Conservative: 246
 Best Local Similarity: 54.31% Mismatches: 264
 Query Match: 57.18% Indels: 63
 DB: 1 Gaps: 10

US-09-873-409-5 (1-1222) x US-08-181-471-2 (1-4646)

QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
 DB 575 ATGGTGGTGGAACTTGGCTGCCATCATCCATGGGGCTGGACTTCTCTCATGATGCTG 634
 QY 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
 DB 635 GTGTTTGGAGAAATGACAGATATCTTTGCAAAATGACAGAAATTTAGAAAGATCTGATGTC 694
 QY 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
 DB 695 AACATCACTAATAGAGTATGATCAATCAATACAGGGTGTCTTCATGAATCTGGAGGAAGAC 754
 QY 44 -----ArgLeuThrLeuTyrValGlyIleGlyValAlaLeuIlePheGlyTyr 61
 DB 755 ATGACAGGATATGCTATTATTACAGTGGAAATGGTCTGGGGTGTGGTGTGCTGTAC 814
 QY 62 IleGlnIleSerLeuThrIleLeuAlaAArgGlnThrLysArgIleAArgLysGln 81
 DB 815 ATTCAGGTTTCATTTGGTCTGGCAGCTCGAAGACAAAATACACAAAATTAGAAAACAG 874
 QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
 DB 875 TTTTTCATGCTATATGCGACAGAGATAGGCTGGTGTGATGTCACGATGTTGGGAG 934
 QY 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
 DB 935 CTTAACACCCGACTTACAGATGATGCTCTTAAGATTAAATGAAGTTATTGTGACAAAATT 994
 QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
 DB 995 GGAATGTTCTTCAGTCAATGGCAACATTTTCACTGGGTTTATAGTAGGATTACACGT 1054
 QY 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
 DB 1055 GGTGGAAAGCTAACCCCTGTGATTTTGGCCATCAGTCTGCTTTCAGTGTGCTGCTGCT 1114
 QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
 DB 1115 GTCTGGCAAGATACATCTTCAATTTACTGTATTAAGAACTCTTAGCGTATGCAAAAGCT 1174
 QY 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
 DB 1175 GGACAGTAGCTGAAGAGGCTTGGCAGCAATTAGAACTGTGATGCAATTGGAGGACAA 1234
 QY 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
 DB 1235 AAGAAAGAACTTGAAA----- 1249
 QY 221 ProGlnTrpLeuLeuSerCysValLeu**PheValArgTyrThrGlnAsnLeuLysAsp 240

DB 1250 -----AGGTACAAACAAAAATTTAGAAAGAA 1273
 QY 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
 DB 1274 GCTAAAGAAATTTGGGATAAAGAAAGCTATTACAGCAATATTCTATAGGTGCTGCTTTC 1333
 QY 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
 DB 1334 CTGCTGATCATCATCTTATGCTCTGGCCTTCTGGTATGGACCACTTGTGCTCTCTCA 1393
 QY 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
 DB 1394 GGGGAA-----TATTCTATTGGACAAAGTACTGCTGTTCTTCTGCTATTAAATTTGG 1447
 QY 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
 DB 1448 GCTTTTAGTGTGGACAGCATCTCCAAGCATTTGAAGCAATTTGCAAAATCAGAGAGGCA 1507
 QY 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
 DB 1508 GCTTATGAATCTTCAAGATATTTGATATAGCCCAAGTATTGACAGCTATTTCGAAGAGT 1567
 QY 341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
 DB 1568 GGGCACAAACAGATAATATTAAAGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 1627
 QY 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
 DB 1628 CCATCTCGAAAAAGAAATTAAGATCTTGAAGGCCCTGAACCTGAAGGTGCAGAGTGGCAG 1687
 QY 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
 DB 1688 ACGTGGCCCTGTTGGAAACAGTGGCTGTGGAGAGACCAACAGTCCAGCTGATGTCAG 1747
 QY 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
 DB 1748 AGGCTCTATGACCCACAGAGGGGATGGTCAGTGTGTGATGGACAGGATATTAGGACCATA 1807
 QY 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
 DB 1808 AATGTAAAGTTTCTACGGGAAATCATTTGTTGGTGTGAGTCAGGAACCTGTATTGTTGCC 1867
 QY 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
 DB 1868 ACCACATAGCTGAAACCAATTCGCTATGCCGTGAAATGTCAACCATGATCAGATTGAG 1927
 QY 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
 DB 1928 AAGAGTGTCAAGGAAGCAATGCTATGACTTTATCATGAAACTGCCTCATAAATTTGAC 1987
 QY 481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
 DB 1988 ACCCTGTTTGGAGAGAGAGGGGCCCAAGTTGAGTGTGGGCAAGAGCAGAGGATCGCCATT 2047
 QY 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
 DB 2048 GCACGTGCCCTGTTCCCAACCCCAAGATCTCTGCTGTGATGAGGCCACGTCAGGCTTG 2107
 QY 521 AspSerGluSerLysSerAlaValGlnAlaLeuGluLysAlaSerLysGlyArgThr 540
 DB 2108 GACACAGAAAGCGAAGCAGTGGTTTCAAGTGGCTCTCGATAAGGCCAGAAAAGGTCCGACC 2167
 QY 541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560
 DB 2168 ACCATTGTGATGACTCATCTGTTGTCTACAGTTTCGTAATGTGCTGACATCGCTGCTTTC 2227
 QY 561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580
 DB 2228 GATGATGGAGTCACTTGTGGAGAAAGGAATCATGTAATCATGTAAGAGAGAAAGGCATT 2287
 QY 581 TyrTyrSerLeuValMetSerGln-----AspIleLysLysAlaAsp 594

COUNTRY: U.S.A.
 ZIP: 20005-5701
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/793,610
 FILING DATE: 07-MAR-1997
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: DE P 44 31 973.8
 FILING DATE: 08-SEP-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: DE 195 03 952.1
 FILING DATE: 07-FEB-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/EP95/03175
 FILING DATE: 10-AUG-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Berman, Richard J.
 REGISTRATION NUMBER: 39,105
 REFERENCE/DOCKET NUMBER: P1614-7007
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 638-5000
 TELEFAX: (202) 638-4810
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 6505 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: circular
 MOLECULE TYPE: DNA
 US-08-793-610-5

Alignment Scores:

Pred. No.: 0 Length: 6505
 Score: 3521.50 Matches: 681
 Percent Similarity: 73.92% Conservative: 246
 Best Local Similarity: 54.31% Mismatches: 264
 Query Match: 59.18% Indels: 63
 DB: 2 Gaps: 10

US-09-873-409-5 (1-1222) x US-08-793-610-5 (1-6505)

QY	121	AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys	140
DB	2387	GGAAATGCTTTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGGATTTACACGT	2446
QY	141	GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla	160
DB	2447	GGTTGGAAGCTAACCCCTTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTGAGCTGT	2506
QY	161	AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTrpSerLysAla	180
DB	2507	GTCGGGCAAGATACATCTCTTCTATCTGATAAAGAACTCTTAGCGTATGCAAAAGCT	2566
QY	181	GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln	200
DB	2567	GGAGCAGTAGCTGAAGAGCTCTGGCAGCAATTAGAATCTGTGATTCATTTGGAGACAA	2626
QY	201	GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe	220
DB	2627	AGAAAGAACTTGAA-----	2641
QY	221	ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp	240
DB	2642	-----AGGTACAAACAAATTTAGAGAA	2665
QY	241	AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr	260
DB	2666	GCTAAAGAAATTTGGGATAAAGAAAGCTATTACAGCAATATTTCTATAGTGTGCTTTC	2725
QY	261	PhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrTrpTyrGlyThrSerLeuIleLeuAsn	280
DB	2726	CTGCTGATCTATGCACTTATGCTCTGGCTTCTGTTGATGGACCACTTGTGTCTCTCA	2785
QY	281	GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer	300
DB	2786	GGGAA-----TATCTATTGGCAAGTACTCCTGATTTCTTTCTGTATTAAATGGG	2839
QY	301	SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla	320
DB	2840	GCTTTTAGTGTGGACAGCATCTCCAGCAATGAAGCATTTGCAATGCAGAGAGCA	2899
QY	321	AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla	340
DB	2900	GCTTATGAAATCTTCAAGATAATTTGATAATAAGCCAAATTTGACAGACTATTCCAAAGAT	2959
QY	341	GlyTyrLysProGluSerIleGlyThrValGluPheLysAsnValSerPheAsnTyr	360
DB	2960	GGCACAACCAAGATATAATTAAGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTAC	3019
QY	361	ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu	380
DB	3020	CCATCTCGAAAGAAAGTTAAGATCTTGAAGGGCTGAACCTGAAGGTGCAGAGTGGGCAG	3079
QY	381	ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln	400
DB	3080	ACGGTGGCCCTGTTGGAAACAGTGGCTGTGGAGAGAGCACACAGTCCAGCTGATGCAG	3139
QY	401	ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu	420
DB	3140	AGGCTCTAGACCCACAGAGGGGATGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	3199
QY	421	AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly	440
DB	3200	ATGTAAGGTTTCTACGGGAATCATTTGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	3259
QY	441	ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluMetGlu	460
DB	3260	ACCACGATAGCTGAAACATTCGCTATGCGTGTGAAATGTCCACCATGGATGAGATTGAG	3319
QY	461	ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn	480
DB	3320	AAAGCTGTCAAGGAAGCCCAATGCCTATGACTTTATCATGAAATGCTCATAAATTGAC	3379
QY	481	ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle	500

[illegible]

Qy	1207	ArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer	1220
		::: ::: :::	
Db	5600	GCACAGAAAGGCATCTATTTTCAATGGTCAGTGTCCAGGCT	5641

```

RESULT 3
US-08-793-610-6
; Sequence 6, Application US/08793610
; Patent No. 5858744
;
; GENERAL INFORMATION:
;
; APPLICANT: BAUM, Christopher
; APPLICANT: STOCKING-HARBERS, Carol
; APPLICANT: OSTERTAG, Wolfram
;
; TITLE OF INVENTION: RETROVIRAL VECTOR HYBRIDS AND THE USE THEREOF
; TITLE OF INVENTION: FOR GENE TRANSFER
; NUMBER OF SEQUENCES: 6

```

Alignment Scores:		
Pred. No.:	0	Length:
Score:	3521.50	Matches:
Percent Similarity:	73.9%	Conservative:
Best Local Similarity:	54.31%	Mismatches:
Query Match:	57.18%	Indels:
DB:	2	Gaps:
		9318

US-09-873-409-5 (1-1222) x US-08-793-610-6 (1-9318)

Qy	1	MecLeuLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu	20
		: : :	
		: : :	
		: : :	
Db	1926	ATGGTGGTGGAACTTTGGCTGCCATCCATCGGGGTGGACTTCCTCTCATGATCGT	1985
Qy	21	ValLeuGlyGluMetSerAsp-----AsnLeuIleSer	31
Db	1886	GTGTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAATATTAGAAGATCTCATGTCA	2045

QY 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuGln 400
DB 3039 ACGTGGCCCTGCTGGTGAACACATGGCTGCTGGGAAGACAGACACACAGTCCAGCTGATGCAG 3098
QY 401 ArgLeuTyrAspProAspGlyPheIleMetValAspGluAsnAspIleAlaAlaLeu 420
DB 3099 AGGCTCTATGACCCACAGAGGGGATGGTCAGTGTGTGATGGACAGGATATTAGGACCAATA 3158
QY 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
DB 3159 AATGTAAGGTTTCTACGGGAATCATTTGTTGGTGGTGCAGTACGGAACCTGTTGTTGCC 3218
QY 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluMetGlu 460
DB 3219 ACCACGATGCTGAACACATTCGCTATGGCCGCTGAAATGTCCACCATGATGAGATTGAG 3278
QY 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
DB 3279 AAAGCTGTCAAGGAAGCCATGCTATGACTTTATCATGAACTGCTCATTAATTTTGAC 3338
QY 481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
DB 3339 ACCCTGGTTGGAGAGAGAGGGCCCAAGTTGAGTGGTGGCAGAGACAGAGGATCGCCATT 3398
QY 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
DB 3399 GCACGTGCCCTGCTGCACACCCCAAGATCCTCTCTGGATGAGGCCACAGTCCAGCTTG 3458
QY 521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThr 540
DB 3459 GACACGAAGCGAAGCATGGTTTCAGGTGGCTCTCGATAAGCCAGAAAGGTCGGACC 3518
QY 541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560
DB 3519 ACCATTGCTAGTACATCGTTGCTGTACAGTTCGTAATGCTGACGTCATCGCTGTTTC 3578
QY 561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580
DB 3579 GATGATGGAGTCATTGTGGAGAAAGGAATCATGATGAACATCATGAAGAGAAAGCAATT 3638
QY 581 TyrTyrSerLeuValMetSerGln-----LysSerIleLys-----AspIleLysLysAlaAsp 594
DB 3639 TACTTCAAACTTGTCAATGCAGACAGCAGGAAATGAAAGTTGAATTGAAATGCAAGCT 3698
QY 595 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis 614
DB 3699 GATGAATCCAAAGTGAATGATGCTCTGGAAATGCTCTCAATGATTCATGATCCAGT 3758
QY 615 SerVal-----LysSerIleLys-----LysSerIleLys-----SerAspPheIleAspLysAla 627
DB 3759 CTAATAAGAAAAAGATCAACTCGTAGGAGTGTCCGTGGATCAACAGCCCAAGACAGAAAG 3818
QY 628 GluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeu 647
DB 3819 CTTAGTACCACAAAGAGGCTCTGGATGAAGATATACCTCCAGTTTCTTTGGAGGATATG 3878
QY 648 LysLeuAsnLysProGluTyrProPheValValLeuGlyThrLeuAlaSerValLeuAsn 667
DB 3879 AAGCTAAATTAACCTGAATGGCTTATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3938
QY 668 GlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn 687
DB 3939 CGAGGCTGCAACCAAGCATTTGCAATAATATTTTCAAGATATATAGGGGTTTTTACAAGA 3998
QY 688 ---AsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIle 706
DB 3999 ATTGATGATCCTGAAACAAACACAGAGATAGTAACTGTTTCTACTATTGTTCTTAGCC 4058
QY 707 LeuGlyValIleCysValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGly 726
DB 4059 CTTGGAATATTCTTTTATACATTTTCTTCAGGGTTTCATATTGCAAAAGCTGGA 4118
QY 727 GluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIle 746

DB 4119 GAGATCCCTCACCAACGGCTCCGATACATGTTTCCGATCCATGCTCAGACAGGATGG 4178
QY 747 AlaTyrPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAsp 766
DB 4179 AGTTGGTTGATGACCTTAAACACACACTGGAGCATTTGACTACAGGCTCCGCAATGAT 4238
QY 767 IleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThr 786
DB 4239 GCTGTCTCAAGTTAAAGGGCTATAGTTTCCAGGCTGCTGTAATTTACCAGAATATAGCA 4298
QY 787 AsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyrGluMetThrPheLeuIle 806
DB 4299 AATCTGGGACAGGAATAATATATCTTATCTATGTTGGCAACTAAACACTGTTTACTC 4358
QY 807 LeuSerIleAlaProValLeuAlaThrGlyMetIleGluThrAlaAlaMetThrGly 826
DB 4359 TTAGCAATTTGATCCCATCATTCATGATGAGGAGTGTGTGAATGAAATGTTGTCTGA 4418
QY 827 PheAlaAsnLysAspLysGlnLeuLysHisAlaGlyLysIleAlaThrGluAlaLeu 846
DB 4419 CAAGCACTGAAAGATAAGAAAGAACTAGAAGGTGCTGGGAAGATCGTACTGAAGCAATA 4478
QY 847 GluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGlu 866
DB 4479 GAAAACTTCCGAACCGTTGTTCTTTGACTCAGGAGCAGAGCTTTGAACATATGATGCT 4538
QY 867 GluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCys 886
DB 4539 CAGAGTTTGCAGGTACCATACAGAAACTCTTTGAGAAAGACACACATCTTTGGAATTACA 4598
QY 887 TyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAla 906
DB 4599 TTTTCTCTTCCACCGCAATGATGATTTTCTCTATGCTGGATGTTTCGGTTTGGAGCC 4658
QY 907 TyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIle 926
DB 4659 TACTTGGTGGCACAATAAATCATGAGCTTTGAGGATGTTCTCTGTAGTATTTTTCAGCTGT 4718
QY 927 AlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAla 946
DB 4719 GTCTTTGGTGCATGGCTGGGGCAAGTCACTTCATTTGCTCTCCGACTATGCCAAGCC 4778
QY 947 LysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysPheAsnIleAspSerArg 966
DB 4779 AAAATATCAGCAGCCACATCATCATTCATTTGAAAAACCCCTTTGATTGACACTAC 4838
QY 967 SerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSer 986
DB 4839 AGCAGCGAAGGCTTAATGCCGAACACATTTGGAAGGAAATGTCCACATTTTGGTGAAGTTGA 4898
QY 987 PhePheTyrProCysArgProAspValPheIleLeuArgGlyLysSerLeuSerIleGlu 1006
DB 4899 TTCAACTATCCACCGACCGACATCCAGTCTTCAGGACCTGAGCTGGAGGTGAAG 4958
QY 1007 ArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGln 1026
DB 4959 AAGSGCCACAGCCTGCTCTGTGGGCGACAGTGGCTGTGGGAAGACACAGTGGTCCAG 5018
QY 1027 LeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAla 1046
DB 5019 CTCTGGAGCGGTTCTACGACCCCTTGGCAGGGAAGTGTCTGTGATGGCAAGAAATA 5078
QY 1047 LysGluLeuAsnValGlnTyrLeuArgSerGlnIleAlaIleValProGlnGluProVal 1066
DB 5079 AAGCAGCTGAATTTTCAGTGGCTCCGAGCACACCTGGGCGCATCTGTCCGAGGACCCATC 5138
QY 1067 LeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValPro 1086
DB 5139 CTGTTTACTGAGCATTTCTGAGAACATTTGCTATGGAGACACACAGCCGGTGTGTCA 5198
QY 1087 LeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeu 1106

Db 5199 CAGGAAGAGATCGTGGGCGAGCAAGAGGCGCAACATACATCCCTTCATCGAGTCACTG 5258
QY 1107 ProGluysrYrAenThrGlnValGlyLeuLysGlnValGlnLys 1126
Db 5259 CCTAATAATATAGCTAAAGTAGGAGACAAAGGAAGTCTCTGGTGCCAGAAA 5318
QY 1127 GlnArgLeuAlaIleAlaArgAlaLeuLysGlnLysProLysIleLeuLeuLeuAspGlu 1146
Db 5319 CAAGCATTCGCATAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5378
QY 1147 AlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAla 1166
Db 5379 GCCACGTCAGCTCGGATACAGAAAGTGAAGAGTTGTCCAGAGCCCTGGCAAGGCC 5438
QY 1167 ArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAsp 1186
Db 5439 AGAAGAGCCGCGACCTGTCATGTGATTGCTCCACCGCTGTCCACCATCCAGAAATGAGAC 5498
QY 1187 LeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeu 1206
Db 5499 TTAATAGTGTGTTTCAGATGCGCAGAGTCAAGAGCATGCGCAGCATCAGCAGCTGCTG 5558
QY 1207 ArgAsnArgAspIleYrPheLysLeuValAsnAlaGlnSer 1220
Db 5559 GCACAGAAAGGCATCTATTTTCAATGGTCAGTGTCCAGGCT 5600
RESULT 4
US-08-784-649A-1
; Sequence 1, Application US/08784649A
; Patent No. 5830697
; GENERAL INFORMATION:
; APPLICANT: Sikic, Branimir I
; APPLICANT: Chen, Gang
; TITLE OF INVENTION: P-GLYCOPROTEIN MUTANT RESISTANT TO
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Fish & Richardson
; STREET: 2200 Sand Hill Road
; CITY: Menlo Park
; STATE: CA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/784,649A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: Reg.No. 5830697 36,677
; REFERENCE/DOCKET NUMBER: 06037/007001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-322-5070
; TELEFAX: 415-854-0875
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4264 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-784-649A-1
Alignment Scores:
Pred. No.: 0 Length: 4264
Score: 3520.00 Matches: 682
Percent Similarity: 74.00% Conservative: 246
Best Local Similarity: 54.39% Mismatches: 262

Query Match: 57.15% Indels: 64
DB: 2 Gaps: 11
US-09-873-409-5 (1-1222) x US-08-784-649A-1 (1-4264)
QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 292 ATGGTGGTGGGAACCTTGGCTGCCATCATCCATCGGGCTGGACCTTCCTCTCATGATGCTG 351
QY 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
Db 352 GTGTTTGGAGAATGACAGATATCTTTGCAATGCGAGAAATTTAGAGATCTGATGTCA 411
QY 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
Db 412 AACATCACTAATAGAGTGTATCATGATACAGGTTCTTCATGAATCTGGAGGAAGAC 471
QY 44 -----ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db 472 ATGACCAGGTATGCTTATTATACAGTGGAAATGGTCTGGGTGCTGGTGTGCTGTAC 531
QY 62 IleGlnIleSerLeuTyrIleThrAlaAlaArgGlnThrIleArgIleArgLysGln 81
Db 532 ATTCAGGTTTCATTTTGGTGCCTGCGCAGCTGGAAGCAAAATACACAAATTTAGAAACAG 591
QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101
Db 592 TTTTTCATGCTAATAGCAGAGGATAGCTGGTTGATGTCGACCATGTTGGGGAG 651
QY 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db 652 CTTAACCCCGACTTACAGATGATGCTCCAAAGATTAAATGAAGGAATTTGTCACAAAAT 711
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db 712 GGA---ATGTTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGAATTCACGT 768
QY 141 GlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db 769 GGTGGAGACTAACCTTGTGATTGGCCATCAGTCTCTGTTCTTGGACTGTGACGTGCT 828
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db 829 GTCTGGGCAAGATACATCTCTCAATTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 888
QY 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db 889 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTTAGAACTGTGATTGCAATTTGGAGGACAA 948
QY 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTyrPheTyrPhe 220
Db 949 AAGAAGAACTTGAA----- 963
QY 221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
Db 964 -----AGGTACAAACAAAAATTTAGAAAGAA 987
QY 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
Db 988 GCTAAAGAAATTTGGGATNAAGAAAGCTATTACAGCAATATTCTATAGTGTGCTTTC 1047
QY 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsn 280
Db 1048 CTGCTCATCTATGATCTTATGCTCTGGCTTCTGGTATGGGACCACTTGGTCTCTCA 1107
QY 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
Db 1108 GGGGAA-----TATTCATTTGGCAAGTACTCAGCTGATTTCTTTCTGTTAATTAATGGG 1161
QY 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaGlyAla 320
Db 1162 GCTTTTAGTGTGGACAGGCATCTCCAGCATTTGAAGCATTTGCAAAATGCAAGAGGACGA 1221

Qy	321	AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla	340
Db	1222	GCATTAGAAATCTTCAAGATAATTTGATAAAGCAAGTATTGACAGCTATTTCGAAGAT	1281
Qy	341	GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr	360
Db	1282	GGGCACAAACCGATATATTAAAGGAAATTTGGAAATTCAGAAATGTTCACCTTCAGTTAC	1341
Qy	361	ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu	380
Db	1342	CCATCTCGAAAGAAGTTAAGATCTTGAAGGCCCTGAACCTGAAGTGCAGAGTGGGCAG	1401
Qy	381	ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln	400
Db	1402	ACGGTGGCCCTGGTTGGAAACAGTGGCTGTGGGAAGAGCACCAAGCTCCAGCTGATGCAG	1461
Qy	401	ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu	420
Db	1462	AGGCTCTATACCCACAGAGGGGATGGTCAGGTGTGATGGACAGGATATTAGGACCATA	1521
Qy	421	AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly	440
Db	1522	ATGTAGAGTTTCTACGGGNAATCATTTGGTGTGTGGTGCAGCAACCTGTATTGTTGCC	1581
Qy	441	ThrThrIleSerAsnIleTyrGlyArgAspAspValThrAspGluGluMetGlu	460
Db	1582	ACCACGATAGCTGAAACATTTGCTATGCGCTGAAATGTCAACATGGATGAGATGAG	1641
Qy	461	ArgAlaalaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn	480
Db	1642	AAAGCTGTCGAAGAAGCCATGCGCTATGACTTTATCATGAAATCGCTCATAAATTTGAC	1701
Qy	481	ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle	500
Db	1702	ACCTCGTTGGAGNAGAGGGGCCAGTTGATGTTGGTCGACAGCAGAGATGCCATT	1761
Qy	501	AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu	520
Db	1762	GCACGTGCCCTGGTTGCAACCCCAAGATCTCTCTGTGATGAGGCCACGCTCAGCCTTG	1821
Qy	521	AspSerGluSerLysSerAlaValGlnAlaLeuGluLysAlaSerLysGlyArgThr	540
Db	1822	GACACAGAAGCGAAGCAGCTGGTTTCAGGTGGCTCTGGATTAAGGCCAGAAAAGTGGACC	1881
Qy	541	ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu	560
Db	1882	ACCATGTGTAGTACTATCGTTTGTCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTC	1941
Qy	561	LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu	580
Db	1942	GATCATGGATCATTTGGGAGAAGGAATCATGATGAACCTCATCAAGAGAAGAGGCATT	2001
Qy	581	TyrTyrSerLeuValMetSerGln-----AspIleLysLysAlaAsp	594
Db	2002	TACTTCAAACTTGTCAATATGCACAGCAGAGGAAATGAAGTTGAATTAGAAAATGCAGCT	2061
Qy	595	GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis	614
Db	2062	GATGAATCCAAAAGTGAATTTGATCGTGGAAATGTCTTCAATGATTCAGATCCAGT	2121
Qy	615	SerVal-----LysSerIleLys---SerAspPheIleAspLysAla	627
Db	2122	CTAATAGAAAAGATCAACTCGTAGGAGTCCGTGGATCACAAGCCCAAGACAGAAG	2181
Qy	628	GluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuIleLeu	647
Db	2182	CTTAGTACCAAAAGAGGCTCTGGATGAAGAGTATACCTCCAGTTTCCTTTTGGAGGATTATG	2241
Qy	648	LysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsn	667
Db	2242	AAGCTAAATTTAACTGAATGGCCCTTATTTGTTGGTGGTATTTTGTGGCATATATAAT	2301
Qy	668	GlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn	687

Db	2302	GGAGGCGCTGCAACAGCAATTCGAATAATATTTTCAAAAGATATTAGGGGTTTTTCAAGA	2361
Qy	688	--AsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIle	706
Db	2362	ATTGATGATCTCGTAACAAACAAACAGAGAAATAGTAAGTCTGTTTTCATATTTGTTCTAGCC	2421
Qy	707	LeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGly	726
Db	2422	CTTGGAATTATTTCTTTTATTACATTTTTTCTTCAGGGTTTTCACATTTTGGCAAGCTGGA	2481
Qy	727	GluIleLeuThrMetArgLeuArgHisIleuAlaPheLysAlaMetLeuTyrGlnAspIle	746
Db	2482	GAGATCTCTACCAAGCGGCTCGATACATGTTTTCGATCCATGCTCAGACAGAGATGTG	2541
Qy	747	AlaTrpPheAspGlnLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAsp	766
Db	2542	AGTTGGTTTGATACCCCTAAAACACCACTGGAGCAATTTGACTCACCAGGCTCGCCAAATGAT	2601
Qy	767	IleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThr	786
Db	2602	GCTGCTCAAGTTAAAGGGGCTATAGTTCACGCTTGCTGTAAATTTACCCAGAAATATAGCA	2661
Qy	787	AsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIle	806
Db	2662	AATCTTGGGACAGAAATAATATATCTTCATCTATGTTGGCAACTAAACACTGCTTACTC	2721
Qy	807	LeuSerIleAlaProValIleuAlaValThrGlyMetIleGluThrAlaAlaMetThrGly	826
Db	2722	TTAGCAATTTGACCAATCATTCATAGCAGAGAGTGTGTAAATGAAATGTTGCTCTGGA	2781
Qy	827	PheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeu	846
Db	2782	CAAGCACTGAAAGATAAAGAAAGCACTAGAAAGTGTCTGGGAAGATCGCTACTGAAGCAATA	2841
Qy	847	GluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGlu	866
Db	2842	GAATACTTCCGAACCGTTGTTCTTTGACTCAGGAGCAGAAAGTTTGAACATATGTATGCT	2901
Qy	867	GluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCys	886
Db	2902	CAGAGTTTGCAGGTACCATCAGAAACCTCTTTGAGGAAGCAGACACATCTTTGGAAATTACA	2961
Qy	887	TyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAla	906
Db	2962	TTTTCTTTCACCCAGGCAATGATGATATTTTCTATGCTGGATGTTTCCGGTTTGGAGCC	3021
Qy	907	TyrIleuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIle	926
Db	3022	TACTTGTGGCACATAAACTCATGAGCTTTGAGGATGTTCTGTAGTATTTTTCAGCTGTT	3081
Qy	927	AlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAla	946
Db	3082	GTCTTTGTTGGTCCCATGCGCTGGGCAAGTCAGTTCATTTGCTCTGCTCATATGCCAAAGCC	3141
Qy	947	LysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArg	966
Db	3142	AAATATACAGACCCCATCATCATGATCATTTGAAATAAAACCCCTTTGATTTGACAGCTAC	3201
Qy	967	SerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSer	986
Db	3202	AGCACGGAAGCGCTAATGTCGGAACACATTTGGAAGGAAATGTACATTTTGGTGAAGTTGTA	3261
Qy	987	PhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGlu	1006
Db	3262	TTCAACTATATCCCAACCCGACCGACATCCCAAGTCTTTCAGGACATGAGCGCTCGAGGCTGAAG	3321
Qy	1007	ArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGln	1026
Db	3322	AAGGCCACAGACGTGTGCTCTGTGTGGCAGCAGTGTGCTGTGGAGAGACACAGTGGTCCAG	3381
Qy	1027	LeuLeuGlnArgLeuTyrAspProValGlnGlnValLeuPheAspGlyValAspAla	1046

Db 3382 CTCCTGAGCGGTTCTACGACCCCTTGGCAGGGAAGTGTCTGATGCGCAAGAAATA 3441
QY 1047 LysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProVal 1066
Db 3442 AAGCGACTGAATGTTGAGTGGCTCCGAGACACCTGGGATGCTGTCCAGGAGGCCATC 3501
QY 1067 LeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValPro 1086
Db 3502 CTGTTTGAATGTCAGCATTCGTGAGAACATTCCTATGGAGACAACAGCCGGTGGTGCA 3561
QY 1087 LeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeu 1106
Db 3562 CAGGAAGAGATTGTGAGGCGCAGCAAGAGGAGCCCAACATACATGCTTCATCGAGTCACTG 3621
QY 1107 ProGluLysTrpAsnThrGlnValGlyLeuLysGluAlaGlnLeuSerGlyGlyGlnLys 1126
Db 3622 CCTAATTAATATAGCACTAAGTAGGAGACAAAGGAACCTCAGCTCTCTGGTGGCCAGAAA 3681
QY 1127 GlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuAspGlu 1146
Db 3682 CRAAGCATTCGCATAGCTGCTGCGCTTGTAGACAGCCTCATATTTTGTCTTTGGATGAA 3741
QY 1147 AlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAla 1166
Db 3742 GCCACGTCAGCTCTGGATACAGAAAGTGAAGGTTGTCCAGAAGCCCTGGCAAGGCC 3801
QY 1167 ArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAsp 1186
Db 3802 AGAAGAGCGCGCACCTGTCATGTGATGTTCTCACCGCTGTCCACCATCCAGATGCGAGC 3861
QY 1187 LeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeu 1206
Db 3862 TTAATAGTGTGTTTCAGATGTCAGAGTCAAGGAGTCAAGGAGTGGCAGCATCAGCAGCTGCTG 3921
QY 1207 ArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
Db 3922 GCACAGAAAGGCATATTTTCAATGGTTCAGTGTCCAGGCT 3963

RESULT 5

US-08-784-649A-5
Sequence 5, Application US/08784649A
Patent No. 5830697
GENERAL INFORMATION:
APPLICANT: Sikic, Branimir I
APPLICANT: Chen, Gang
TITLE OF INVENTION: P-GLYCOPROTEIN MUTANT RESISTANT TO
CYCLOSPORIN MODULATION
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESS: Fish & Richardson
STREET: 2200 Sand Hill Road
CITY: Menlo Park
STATE: CA
COUNTRY: USA
ZIP: 94025
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/784,649A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sherwood, Pamela J
REGISTRATION NUMBER: Reg.No. 5830697 36,677
REFERENCE/DOCKET NUMBER: 06037/007001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-322-5070
TELEFAX: 415-854-0875
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:

LENGTH: 4264 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-784-649A-5
Alignment Scores:
Pred. No.: 0 Length: 4264
Score: 3520.00 Matches: 682
Percent Similarity: 74.00% Conservative: 246
Best Local Similarity: 54.39% Mismatches: 262
Query Match: 57.15% Indels: 64
DB: 2 Gaps: 11
US-09-873-409-5 (1-1222) x US-08-784-649A-5 (1-4264)
QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 292 ATGGTGGTGGGAACCTTGGCTGCCATCATCCATGGGCTGGACCTTCCTCTCATGATGCTG 351
QY 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
Db 352 GGTGTTGGAGAAATGACAGATATCTTGCAAATGACGAGAAATTAGAACATCTGATGTCA 411
QY 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
Db 412 AACATCCTCACTAATAGAAGTGATCAATGATACAGGGTCTCTCATGAATCTGGAGGAAGAC 471
QY 44 -----ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db 472 ATGACCCAGTATGCTTATTATACAGTGGAAATTTGGTGGGGTGTCTGCTGCTTAC 531
QY 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db 532 ATTCAGTTTCATTTTGGTGCTGCGCAGCTGGAAGCAAAATACACAAAATTAGAAAACAG 591
QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101
Db 592 TTTTTCATGCTAATAATGCGACAGGAGATAGCTGGTTGATGTGCACGATGTTGGGGAG 651
QY 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db 652 CTTAACACCCGACTACAGATGATGTCTCCAAAGATTAAAGAGGAATTTGGTGCAAAATT 711
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db 712 GGA---ATGTTCCAGTCAATGGCAACATTTTTCACGGGTTTATAGTAGATTACACGT 768
QY 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db 769 GGTGGAAGCTAACCCCTGTGATTTTGGCCATCAGTCTCTGTTCTTGGACTGTGCTGCTGCT 828
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db 829 GTCTGGGCAAGATACATCTTCACTTACTGATAAAGAACTCTTACCGTATGCAAAAGCT 888
QY 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db 889 GGAGCAGTAGCTGAAGAGTCTTGGCAGCAATTAGAACCTGTGATTCATTTGGAGACAA 948
QY 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
Db 949 AAGAAAGAACTTGAA----- 963
QY 221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
Db 964 -----AGGTACAAACAAAATTTTAGAAGAA 987
QY 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
Db 988 GCTAAAGAAATTGGGATAAGAAAGCTATTACAGCCCAATATTTCTATAGTGTGCTGCTTTC 1047

QY 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheThrTyrGlyThrSerLeuLeuLeuAsn 280
 Db 1048 CTGCTGATCATCTTATGCTTGGCGCTTCTGTATGGACCACTTGGTCTCTCA 1107
 QY 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
 Db 1108 GGGGAA-----TATTCTATTGGACAAGTACTCAGTGTATCTTTCTGTATTAATTGGG 1161
 QY 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
 Db 1162 GCTTTTGGTGGACAGCAATCTCCAGCAATTTGAAGCATTTGCAAAATGCAAGAGGAGCA 1221
 QY 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
 Db 1222 GCTTATGAAATCTCAAGATAATTGATAAAGCCCAAGTATTGACAGCTATTTCGAAGAGT 1281
 QY 341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
 Db 1282 GGGCACAACCAAGATATATTAAAGGGAATTTTGGAAATTCAGAAATGTTCACTTCAGTTAC 1341
 QY 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
 Db 1342 CCATCTCGAAAGAAAGTTAAGATCTTGAAGGGCTGAACCTGAAGTGCAGATGGGCAG 1401
 QY 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuGln 400
 Db 1402 ACGGTGGCCCTGGTTGGAACAGTGGCTGTGGGAAGAGACACACAGTCCAGCTGATGAC 1461
 QY 401 ArgLeuTyrAspProAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
 Db 1462 AGGCTCTATGACCCACAGAGGGGATGGTCAGTGTGTGGAGCAGGATATTAGGACCATA 1521
 QY 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
 Db 1522 ATGTGAAGTTTCTACGGGAATCATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1581
 QY 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspValThrAspGluMetGlu 460
 Db 1582 ACCAGATAGCTGAAACATTCGTATGCGGTGGAATGTCCACCATGATGAGATGAG 1641
 QY 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
 Db 1642 AAAGCTGTCAAGGAAGCCAAATGCTATGATCTTATCATGAACATGCTCATAAATTTGAC 1701
 QY 481 ThrLeuValGlyGlyLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
 Db 1702 ACCCTGGTTGGAGAGAGAGGGCCAGTGTGAGTGTGGGAGAGCAGAGATCGCCATT 1761
 QY 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
 Db 1762 GCAGTGGCCCTGGTTGCAACCCCAAGATCCTCTGCTGGATGAGGCCACGTGAGCCTTG 1821
 QY 521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThr 540
 Db 1822 GACACAGAAGCGAAGCAGTGGTTTCAAGTGGCTCTGGATAAGGCCAAGAAAGTCCGACC 1881
 QY 541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560
 Db 1882 ACCATTGTGATGCTCATCTTTGTCTACAGTTTCGTAATGCTCAGCTCAFCGCTGTTTC 1941
 QY 561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580
 Db 1942 GATGATGGAGTCATTCTGGAGAAAGGAATCATGATGAACTCATGAAGAGAAAGCAATT 2001
 QY 581 TyrTyrSerLeuValMetSerGln-----AspIleLysLysAlaAsp 594
 Db 2002 TACTTCAAACTTGTCACAATGCACAGCAGGAAATGAAGTTGAAATAGAAAATGCAGCT 2061
 QY 595 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis 614
 Db 2062 GATGAATCCAAAAGTAAATGTATGCTTGGAAATGCTTCAATGATTCAGATCCAGT 2121
 QY 615 SerVal-----LysSerIleLys---SerAspPheIleAspLysAla 627

Db 2122 CTAATAAGAAAAAGATCAACTCGTAGGAGTGTCCGTGGATCAAGCCCAACACAGAAAG 2181
 QY 628 GluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeu 647
 Db 2182 CTTAGTACCAAGAGGCTCTGGATGAAGATATACCTCCAGTTTCTTTTGGAGGATATG 2241
 QY 648 LysLeuAsnLysProGluTyrProPheValValLeuGlyThrLeuAlaSerValLeuAsn 667
 Db 2242 AAGCTAAATTTAACTGAATGGCTTATTTTGTGTGTGTATTTTGGCCATTATAAAT 2301
 QY 668 GlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn 687
 Db 2302 GGAGGCTCTCAACCAAGCATTTGCAATAATATTTCAAGATATTAGGGTGTTTTCAAGA 2361
 QY 688 ---AsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIle 706
 Db 2362 ATTGATGATCTGAAACAAACAGCAGCAATAGTAACTTGTTCATATTTCTAGCC 2421
 QY 707 LeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGly 726
 Db 2422 CTTGGAATTTATTTCTTTTATTACATTTTCTTCAGGGTTTCACACTTTGGCAAGCTGA 2481
 QY 727 GluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIle 746
 Db 2482 GAGATCTCTCACAAGCGGCTCCGATACATAGTGTTCGATCCATGCTCAGACAGGATGT 2541
 QY 747 AlaTyrPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAsp 766
 Db 2542 AGTGTGTTTGTAGACCTTAAACACCACTGGAGCATTGACTACCAAGGCTCGCAATGAT 2601
 QY 767 IleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThr 786
 Db 2602 GCTGCTCAAGTTAAAGGGCTATAGTTCAGGCTTGTCTGTAATTACCCAGAATATAGCA 2661
 QY 787 AsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyrGluMetThrPheLeuIle 806
 Db 2662 AATCTTGGGACAGGAATATATATCTTCACTATGTTGGCACTAAACACTGTTTACTC 2721
 QY 807 LeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGly 826
 Db 2722 TTAGCAATTTGTCACCATCATTCGAATGACGAGGATGTTTGAATGAATAATGTTGCTGGA 2781
 QY 827 PheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeu 846
 Db 2782 CAAGCACTGAAAGATAAGAAAGAACTAGAAAGTGTCTGGGAAGATCGCTACTGAAGCAATA 2841
 QY 847 GluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGlu 866
 Db 2842 GAAAACTTCGAAACCGTGTGTTTCTTGACTCAGGAGCAGAAAGTTTGAACATATGTATGT 2901
 QY 867 GluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCys 886
 Db 2902 CAGAGTTTTCAGGTACCATCATCAGAAACTCTTTGAGGAAGACACATCTTTGGAATTACA 2961
 QY 887 TyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAla 906
 Db 2962 TTTTCTCTTCAACCCGCAATGATGATTTTCTATGCTGGATGTTTCCGGTTGGAGCC 3021
 QY 907 TyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIle 926
 Db 3022 TACTTGGTGGCACAATAAACTCATGCTTTGAGGATGTTCTGTGTAGTATTTTCAGCTGT 3081
 QY 927 AlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAla 946
 Db 3082 GTCTTTGGTGCCATGGCCGTGGGGAAGTCAGTTCATTTGCTCCTGACTATGCCAAGCC 3141
 QY 947 LysSerGlyAlaAlaHisLeuPheAlaLeuGluLysLysProAsnIleAspSerArg 966
 Db 3142 AAAATATCAGAGCCCAACATCATCATGATCATTTGAATAAAACCCCTTGTGATGACAGCTAC 3201
 QY 967 SerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSer 986

Db 3202 AGCAGGAGGCGCTAATGCCGAAACACATTTGGAGGAAATGTCCATTTGGTGAAGTTGTA 3261
Qy 987 PhePheTyr-ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGlu 1006
Db 3262 TTCAACTATCCCAACCGACCGACATCCCGAGTCTTCAGGAGCTGAGCCCTGGAGGTTGAAG 3321
Qy 1007 ArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGln 1026
Db 3322 AAGGCCAGAGCTGCTCTGGTGGGCGAGCAGTGGCTGTGGGAAGACACAGTGGTCCAG 3381
Qy 1027 LeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAla 1046
Db 3382 CTCCTGAGCGGTTTACGACCCCTTGGCAGGGAAGTGTCTTGTGTCACAAAGAAATA 3441
Qy 1047 LysGluLeuAsnValGlnTyrLeuArgSerGlnIleAlaIleValProGlnGluProVal 1066
Db 3442 AAGCGACTGAATGTTGAGTGGCTCCGAGCACACCTGGGCATGCTGTCCAGGAGCCCATC 3501
Qy 1067 LeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValPro 1086
Db 3502 CTGTTTGAATGCTGAGCATGCTGAGACATGCTCTATGGAGACACAGCCGGTGTGTCA 3561
Qy 1087 LeuAspGluLeuLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeu 1106
Db 3562 CAGGAAGAGATTGTGAGGCGAGCAAGAGGCGCAACATACATGCTTCATCGAGTCACTG 3621
Qy 1107 ProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyLys 1126
Db 3622 CCTAATAAATATAGCACTAAAGTAGGAGACAAAGGAACTCAGCTCTCTGGTGCCAGAA 3681
Qy 1127 GlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGlu 1146
Db 3682 CAACGATTTGCCATAGCTCTGTCCTTGTGTAGACAGCTCATATTTGCTTTGGATGAA 3741
Qy 1147 AlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAla 1166
Db 3742 GCCACGTCAGCTCTGGATACAGAAAGTGAAGAGTGTGTCCAAAGCCCTGGCAAGGCC 3801
Qy 1167 ArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAsp 1186
Db 3802 AGAAGAGCGCGACCTGCTGATTTGTTGCTCACCCTGTCCACCATCCAGATGCGAGC 3861
Qy 1187 LeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeu 1206
Db 3862 TTATAGTGTGTTTCAGATGCGCAGAGTCAAGAGCATGCGCAGCATCAGCAGCTGCTG 3921
Qy 1207 ArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
Db 3922 GCACAGAAAGGCATCTATTTTCAATGGTCAAGTGTCCAGGCT 3963

RESULT 6

5206352-3
; Patent No. 5206352
; APPLICANT: Roninson, Igor B.; Pastan Ira H.; Gottesman,
; Michael M.
; TITLE OF INVENTION: COMPOSITIONS FOR CLONES CONTAINING DNA
; SEQUENCES ASSOCIATED WITH MULTIDRUG RESISTANCE IN HUMAN CELLS
; NUMBER OF SEQUENCES: 4
; CURRENT APPLICATION NUMBER: US/07/622,836
; FILING DATE: 24-SEP-1990
; PRIOR APPLICATION NUMBER: 892,575
; FILING DATE: 01-AUG-1986
; APPLICATION NUMBER: 845,610
; FILING DATE: 28-MAR-1986
; SEQ ID NO: 3
; LENGTH: 4669
5206352-3
Alignment Scores:
Pred. No.: 0 Length: 4669
Score: 3519.50 Matches: 681

Percent Similarity: 73.84% Conservative: 245
Best Local Similarity: 54.31% Mismatches: 265
Query Match: 57.14% Indels: 63
DB: 6 Gaps: 10
US-09-873-409-5 (1-1222) x 5206352-3 (1-4669)
Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 575 ATGGTGGTGGAACTTTGGCTGCCATCATCCATGGGCTGGACTTCCTCTCATGATGCTG 634
Qy 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
Db 635 GTGTTTGGAGAAATGACAGATATCTTGCATAATGCAGGAAATTTAGAGATCTGATGTCA 694
Qy 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
Db 695 AACATCATCTAATAAGATGATATCAATGATACAGGGTCTTCATGAATCTGGAGGAAGAC 754
Qy 44 -----ArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db 755 ATGACCAAGTATGCTTATTTACAGTGAATTTGGTGGGGTCTGGTGTCTGCTTAC 814
Qy 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db 815 ATTCAAGTTTCAATTTGGTGGCTGGCAGCTGGAAGACAAATACACAAATTAGAAAAACAG 874
Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101
Db 875 TTTTTCATGCTATATAGCAGAGATAGGCTGTTTGTATGTGCACCATGTTGGGGAG 934
Qy 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db 935 CTTAACACCCGACTTACAGATGATGCTCTAAGATTAACTCAAGTTATTTGGTGCACAAAT 994
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db 995 GGAATGCTTTTCAAGTCAATGGCAATTTTCACTGGGTTTATAGTAGGATTTACACGT 1054
Qy 141 GlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db 1055 GGTGGAAGCTAACCTTGTGATTTGGCCATCAGTCCCTGTTCTGGACTGTGACGTGCT 1114
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db 1115 GCTCGGCAAGACTATCTTCACTTACTGATTAAGAACTCTTAGCGTATGCAAAAGCT 1174
Qy 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db 1175 GGAGCAGTAGCTGAAGAGGCTTGGCAGCAATTAGAACTGTGATTCATTGAGGAGCAA 1234
Qy 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
Db 1235 AAGAAAGAACTTGAA----- 1249
Qy 221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
Db 1250 -----AGTACACAAATAATTTAGAGAA 1273
Qy 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
Db 1274 GCTAAAGAAATTTGGGATAAAGAAAGCTATTACAGCAATATTTCTATAGTGTCTGCTTC 1333
Qy 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsn 280
Db 1334 CTGCTGATCTATGATCTTATGCTGGGCTTCTGATGGGACCACTTGGTGTCTCTCA 1393
Qy 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
Db 1394 GGGGAA-----TATCTATTGGACAAAGTACTCAGTGTATTTCTGTATTAAATGGG 1447
Qy 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320


```
QY 1027 LeuLeuGlnArgLeuTyrAspProValGlnGlyValLeuPheAspGlyValAspAla 1046
DB 3668 CTCTGGAGCGGTTCACGACCCCTTGGCAGGGAAGTGTCTTATGCGCAAGAAATA 3727
QY 1047 LysGluLeuAsnValGlnThrLeuArgSerGlnIleAlaIleValProGlnGluProVal 1066
DB 3728 AAGCGACTGAATGTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCCAGGAGCCATC 3787
QY 1067 LeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValPro 1086
DB 3788 CTGTTTACGTCAGCATGCTGAGACATTTGCTTGGAGACAAACAGCCGGGTGTCA 3847
QY 1087 LeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeu 1106
DB 3848 CAGAGAGATCTGTAGGCGAGCAAGAGGCGCCAAATACATGCTCTTCATCGAGTCACTG 3907
QY 1107 ProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLys 1126
DB 3908 CCTAATAAATATAGCACTAAAGTAGGAGACAAAGAACTCAGCTCTCTGTGGCCAGAAA 3967
QY 1127 GlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuAspGlu 1146
DB 3968 CAACGCAATGCCATAGCTGTCGCCCTTTGTAGACAGCTCATATTTGCTTTTGGATGA 4027
QY 1147 AlaThrSerAlaLeuAspAsnSerGluLysValValGlnHisAlaLeuAspLysAla 1166
DB 4028 GCCAGTCAGCTCTGGATACAGAAAGTGAAGAGTTGTCCAAAGCCCTGGACAAAGCC 4087
QY 1167 ArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAsp 1186
DB 4088 AGAAGAGCGCGCACCTGTCATTTGCTTCCAGCCCTGTCCACCATCCAGAGATGCAGAC 4147
QY 1187 LeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeu 1206
DB 4148 TTAATAGTGTGTTTCCAGATGGCAGATCAAGAGATGGCAGCATGCAGCATGCAGCTGCTG 4207
QY 1207 ArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
DB 4208 GCACAGAAAGGCATCTATTTTTCATGGTTCAGTGTCCAGGCT 4249

RESULT 7
US-08-752-447-1
; Sequence 1, Application US/08752447
; Patent No. 5984088
; GENERAL INFORMATION:
; APPLICANT: Mechnetr, Eugene
; APPLICANT: Roninson, Igor B
; TITLE OF INVENTION: Methods and Reagents for Preparing and
; TITLE OF INVENTION: Using Immunological Agents Specific for P-glycoprotein
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boenhen Hulbert & Berghoff Ltd.
; STREET: 300 South Wacker Drive, Seventh Floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/752,447
; FILING DATE: 15-NOV-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5994088nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 95,1121
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-9808
```

```
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4669 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: 1..424
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 425..4264
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: 4265..4669
US-08-752-447-1

Alignment Scores:
Pred. No.: 0 Length: 4669
Score: 3503.50 Matches: 677
Percent Similarity: 73.52% Conservative: 245
Best Local Similarity: 53.99% Mismatches: 269
Query Match: 56.88% Indels: 63
DB: 2 Gaps: 10

US-09-873-409-5 (1-1222) x US-08-752-447-1 (1-4669)

QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
DB 575 ATGGTGTGGGAACCTTGGCTGCCATCATCCATGGGGCTGGACATCTCTCTCATGATGCTG 634
QY 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
DB 635 GTGTTGGGAATAACACATATCTTTGCAATGACGAAATTTAGAACATCTGATGTCA 694
QY 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
DB 695 AACATCACTAATAAGAGTATGATATCAATGATACAGGCTTCTTTCATGAATCTGGAGGAAGAC 754
QY 44 -----ArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuLeuPheGlyTyr 61
DB 755 ATGACCAGGTATGCTTATTATTACAGTGGAAATTTGGTGGGTGCTGGTGTCTGCTTAC 814
QY 62 IleGlnIleSerLeuThrIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
DB 815 ATTCAGGTTTCATTTTGGTGGCTGGCAGCTGGAAGCAAAATACACAAATTAGAAAACAG 874
QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101
DB 875 TTTTTCATGCTATTATGCGACAGGAGATAGGCTGTTTGTATGTGTCACGATGTTGGGAG 934
QY 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspIle 120
DB 935 CTTAACACCCGACTTACAGATGATGTCTCCAAGATTAAAGAAAGAAATTTGGTCAAAAT 994
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
DB 995 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACGTTGTTTATAGTAGATTACACGT 1054
QY 141 GlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
DB 1055 GGTGGGAAGCTAACCTTGTGATTTTGGCCATCAGTCTCTGTTCTTCTGCACTGTGAGTGT 1114
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
DB 1115 GTCTGGGCAAGATACATCTTCTTACTGATAAAGAACTCTTAGCGTAGTGCAGAAAGCT 1174
QY 181 GlyAlaValAlaGluGlnValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
DB 1175 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTCATTTGAGGACAA 1234
QY 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
```


Db	2228	GATGATGGAGTCATTGTCGAGAAAGAAATCATGATGAACATCATGAAGACAGAAAGGCATT	2289
Qy	581	TyTyrSerLeuValMetSerGln-----AspileLysLysAlaAsp	594
Db	2288	TACTTCAAACTTCTGTCACAATGCAGACAGCAGGAAATGAAGTTGAATTAAGAAATGCAGCT	2347
Qy	595	GluClnMetGluSerMetThrTyTyrSerThrGluuArgLysThrAsnSerLeuProLeuHis	614
Db	2348	GATGAATCCAAAAGTCAAAATGATGCTTGGAAATGTCTTCAATGATTTCAAGATCCAGT	2407
Qy	615	SerVal-----LysSerilleLysSerAspPheilleAspLysAlaGlu	628
Db	2408	CTAATAGAAAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAGCCCAACACAGAAAG	2467
Qy	629	GluSerThr---GlnSerLysGluIleSerLeuProGluValSerLeuLysIleLeu	647
Db	2468	CTTAGTACCACAAAGAGGCTCTGGATGAAAGTATACATCTCCAGTTTTCTTTTGGAGGATTATG	2527
Qy	648	LysLeuAsnLysProGluTrpPropheValValLeuGlyThrLeuAlaSerValLeuAsn	667
Db	2528	AAGCTAAATTAATCACTGAATGGGCTTATTTGTGTGGTGTATTTGTGGCCATTATAAAT	2587
Qy	668	GlyThrValHisProValPheSerillePheAlaLysIleIleThrMetPheGlyAsn	687
Db	2588	GGAGGCTCGAACACAGCATTTGCATATATATTTTCAAAGATTTATAGGGGTTTTTACAAGA	2647
Qy	688	---AsnAspLysThrThrLeuLysHisAspAlaGluIleTyTyrSerMetIlePheValIle	706
Db	2648	ATTGATGATCCTGAAACAAACACGACAGAAATAGTAACCTGTTTTTCATCTATGTTTTCTAGCC	2707
Qy	707	LeuGlyValIleCysPheValSerTyTyrPheMetGlnGlyLeuPheTyTyrGlyArgAlaGly	726
Db	2708	CTTGGAAATTAATTTCTTTTATTAATCAATTTTTCTTCAGGGTTTCACATTTTGGCAAGCTGGA	2767
Qy	727	GluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyTyrGlnAspile	746
Db	2768	GAGATCTCTACCAAGCGCTCCGATACATGTTTCCGATCCTCTGACAGGAAGTG	2827
Qy	747	AlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAsp	766
Db	2828	AGTTGGTTTCATGACCCTAAACACCACTGGAGCATTTGACTACCAGGCTCGCCAAATGAT	2887
Qy	767	IleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThr	786
Db	2888	GCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTAAATTTACCAGCAATATAGCA	2947
Qy	787	AsnMetGlyLeuSerValIleIleSerPheIleTyTyrGlyTrpGluMetThrPheLeuIle	806
Db	2948	AATCTTGGGACAGAAATATATATCTTCATCTATGTTGGTGGCACTAACACTGTTACTC	3007
Qy	807	LeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGly	826
Db	3008	TTAGCAATTTGATCCCATCATTTGCAATAGCAGAGATGTTGAAATGAAATGTTTGTCTGGA	3067
Qy	827	PheAlaAsnLysAspLysGlnLeuLysHisAlaGlyLysIleAlaThrGluAlaLeu	846
Db	3068	CAAGCACTGAAAGATAAGAAAGAACTAGAAAGGTGCTGGGAAGATCGCTACTGAAGCAATA	3127
Qy	847	GluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluClnMetTyTyrGlu	866
Db	3128	GAATAACTTCGAAACCGTTGTTTCTTTGACTCGAGGACGAGGTTTGAACATATGTTATGCT	3187
Qy	867	GluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCys	886
Db	3188	CAGAGTTTGGAGGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTTTGGATTACA	3247
Qy	887	TyrAlaPheSerHisAlaPheIleTyTyrPheAlaTyTyrAlaAlaGlyPheArgPheGlyAla	906
Db	3248	TTTTCTTCACCCAGGCAATGATGATATTTTCTATGCTGGATGTTTCCGTTTCGGTTTCGAGCC	3307
Qy	907	TyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIle	926
Db	3308	TACTTGGTGACATAAATCATGAGCTTTGAGGATGTTCTGTAGTATTTTCAGCTGTT	3367


```
QY 927 AlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAla 946
Db 3368 GTCTTTGGTCCATGCGCGGGGCAAGTCAGTTCTTCTGACATGCAAGGCC 3427
QY 947 LysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArg 966
Db 3428 AAAATATCAGCAGCCACATCATCATTCATTTGAAAAACCCCTTTGATTGACAGCTAC 3487
QY 967 SerGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSer 986
Db 3488 AGCAGGAAGGCCTTATGCGGACACATTCGGAAGGAATGTCATTTGTTGAGTTGTA 3547
QY 987 PhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGlu 1006
Db 3548 TTCAACTATCCACCCGACCGACATCCAGTGTCTTCAGGACTGAGCCTGGAGGTGAAG 3607
QY 1007 ArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrServalGln 1026
Db 3608 AAGGCCAGACGCTGCTCTGCTGGGAGCAGTGGCTGTGGGAAGACACAGTGGTCCAG 3667
QY 1027 LeuLeuGluArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAla 1046
Db 3668 CTCTTGAGCGGTTCTACGACCCCTTGGCAGGGGAAGTGTCTTGTATGCCAAGAAATA 3727
QY 1047 LysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProVal 1066
Db 3728 AAGCGACTGATGTTCTAGTGGCTCCGAGCACACCTGGGCATCGTGTCCAGAGCCCATC 3787
QY 1067 LeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValPro 1086
Db 3788 CTGTTTGACTGCAGCATTTGCTGAGACATTTGCTTATGGAGAACACAGCCGGTGTGTCA 3847
QY 1087 LeuAspGluLeuLysGluAlaAlaAsnAlaAsnIleHisSerPheIleGluGlyLeu 1106
Db 3848 CAGGAAGAGATCGTAGGCGGACGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3907
QY 1107 ProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLys 1126
Db 3908 CCTAATATATATAGCATTAAGTAGGAGACAAAGGAATCATGCTCTCTGTGGCCAGAA 3967
QY 1127 GlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuAspGlu 1146
Db 3968 CAACGATTTGCCATAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4027
QY 1147 AlaThrSerAlaLeuAspAsnAspSerGluLysValGlnHisAlaLeuAspLysAla 1166
Db 4028 GCCACGTCAGCTCTGGATACAGAAAGTGAAGAGTTGTCCAAAGGCCCTGGCAAGGCC 4087
QY 1167 ArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAsp 1186
Db 4088 AGAGAAGCGCGACCTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTG 4147
QY 1187 LeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnLeuLeu 1206
Db 4148 TTAATAGTGTGTTTCAAGATGGCAGATCAAGAGGATGCGCAGCATCAGCAGCTGCTG 4207
QY 1207 ArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
Db 4208 GCACAGAAAGGCATCTATTTTCAATGGTCTGAGTGTCCAGGCT 4249

RESULT 8
US-09-316-167-1
; Sequence 1, Application US/09316167
; Patent No. 6365357
; GENERAL INFORMATION:
; APPLICANT: Mechetrner, Eugene
; APPLICANT: Roninson, Igor B
; TITLE OF INVENTION: Methods and Reagents for Preparing and
; TITLE OF INVENTION: Using Immunological Agents Specific for P-glycoprotein
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff Ltd.
```

```
STREET: 300 South Wacker Drive, Seventh Floor
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/316,167
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION NUMBER: 08/752,447
FILING DATE: 15-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: No. 6365357nan, Kevin E
REGISTRATION NUMBER: 35,303
REFERENCE/DOCKET NUMBER: 95,1121
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-913-0001
TELEFAX: 312-913-9808
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4669 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..424
FEATURE:
NAME/KEY: CDS
LOCATION: 425..4264
FEATURE:
NAME/KEY: 3'UTR
LOCATION: 4265..4669
US-09-316-167-1
Alignment Scores:
Pred. No.: 0 Length: 4669
Score: 3503.50 Matches: 677
Percent Similarity: 73.52% Conservative: 245
Best Local Similarity: 53.95% Mismatches: 269
Query Match: 56.88% Indels: 63
DB: 4 Gaps: 10
US-09-873-409-5 (1-1222) x US-09-316-167-1 (1-4669)
QY 1 MetIleGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 575 ATGTGTGGTGGAACTTGTGCTGCCATCATCTCATCTGCGGTGAGTCTCTCTCATGATGCTG 634
QY 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
Db 635 GTGTTGGGAAGATGACATATCTTGGCAATGCAGGAATTTAGAGATCTGATGTCA 694
QY 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
Db 695 AACATCACTAATAGAAAGTATATCAATGATATACAGGCTTCTTCAATGAATCTGGAGGAAGAC 754
QY 44 -----ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db 755 ATGACCAAGGTATGCTTATTATTACAGTGAATTTGGTGGGTGCTGCTGCTGCTTAC 814
QY 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db 815 ATTCAGGTTTCATTTTGGTGGCTGGCAGCTGGNAGACAAATACACAATATAGAAACAG 874
QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
```


875 TTTTTCATGCTATTATGCGACAGGAGATAGGCTGGTTGATGTGCACGATGTGGGAG 934
 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
 935 CTTAAACACCGACTTACAGATGATGCTCCAAGATAATTAAGAGAAATTTGGTGACAAAT 994
 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
 995 GGAATGTTCTTTTCAGTCAATGCGCAACATTTTTCACCTGGGTTTATAGTATTCACAGT 1054
 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
 1055 GGTGTGAAGCTAACCTTGTGATTTGGGCATCAGTCCTGTTCTTGGACTGTGAGCTGCT 1114
 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaIleThrSerLysAla 180
 1115 GTCTGGGCAAGATACCTACTTCTTACTGTATTAAGAACTCTTAGCGTATGCAAAAGCT 1174
 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
 1175 GGAGCAGTAGCTGAAGAGGCTTTGGCAGCAATTAGAACTGTGATTCGATTTGGAGGACAA 1234
 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
 1235 AGAAAGAACTTGAA----- 1249
 221 ProGlnTrpLeuLeuSerCysValLeu**PheValArgTyrThrGlnAsnLeuLysAsp 240
 1250 -----AGGTACAAACAAATTTAGAGAA 1273
 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
 1274 GCTAAAAGAAATTTGGGATAAAGAAAGCTATTACAGCAATATTTCTATAGTGTGCTTTC 1333
 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
 1334 CTGCTGATCTATGATCTTTATGCTGTGGCCTTCTGTATGGGACCCTTGGTCTCTCA 1393
 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
 1394 GGGGAA-----TATTCTATTTGGAAGAACTACTCCTGATTTCTTTCTGTATTAATTGGG 1447
 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
 1448 GCTTTTAGTGTGGACAGGCACTCTCCAAGCATTTGAAGCAITTCGAAATGCAAGAGGACGA 1507
 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
 1508 GCTTATGAAATCTTCAAGATAATTTGATAATAAGCCAAATTTGACAGCTATTGCAAGAGT 1567
 341 GlyTyrLysProGluSerIleGluGlyThrValGluPheIleAsnValSerPheAsnTyr 360
 1568 GGGCAAAACCAAGATAATATTAAAGGAAATTTGGAATTCAGAAATGTTCACITTCAGTTAC 1627
 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
 1628 CCATCTCGAAAGAAAGTAAAGATCTTGAAGGCTTGAACCTGAAGGTGCAGAGTGGGCGAG 1687
 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
 1688 ACGGTGGCCCTGTTGGAAACAGTGGCTGTGGGAAGACACACAGCTCCAGCTGTATGTCAG 1747
 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
 1748 AGGCTCTATGACCCACAGAGGGGATGGTCAAGTGTGTGAGCAGGAGTATATTAGGACCA 1807
 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
 1808 AATGTAAGGTTTCTACGGGAAATCATTTGGTGTGAGTCAGGCACTGTATTTGTTGCC 1867
 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspValThrAspGluMetGlu 460

1868 ACCACGATAGCTGAAAAACATTCGCTATGCGCGTGAAAAATGTCCACATGTGATGAGATTGAG 1927
 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
 1928 AAAGCTGTCAAGGAAGCAATGCTATGATTTTATCATGAAACTGCTCATAAATTTGAC 1987
 481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
 1988 ACCCTGGTTGGAGAGAGAGGGCCAGTTGAGTGTGGTGGCAGAAAGCAGAGGATTCGCATT 2047
 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
 2048 GCACGTGCCCTGTTGCAACCCCAAGATCCTCTGCTGGATGAGCCACAGTCAGGCTTG 2107
 521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThr 540
 2108 GACACAGAAAGCGAAGCAGTGTGTCAGGTGGCTCTGGATAAGCCAGAAAGGTCGGACC 2167
 541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560
 2168 ACCATTGTGATAGCTCATGCTTTTGTCTACAGTTGCTTAATGCTGACGTCTCGTGGTTTC 2227
 561 LysAspGlyMetAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580
 2228 GATGATGGAGTCATTGTGGAGAAAGAAATCATGATGAATCATGAAAGAGAAAGCAAT 2287
 581 TyrTyrSerLeuValMetSerGln-----AspIleLysLysAlaAsp 594
 2288 TACTTCAAACTTGTCAATGCGACAGCAGGAAATGAAGTTGAATTAGAAATGCGACT 2347
 595 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis 614
 2348 GATGAATCCAAAGTCAAAATTCATGCTTTGGAAATGCTCTTCAATGATTCAGATCCAGT 2407
 615 SerVal-----LysSerIleLysSerAspPheIleAspLysAlaGlu 628
 2408 CTAATTAAGAAAAAGATCAACTCGTAGGAGTGTCCGGTGGATCACAAGCCACACAGAAAG 2467
 629 GluSerThr---GlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeu 647
 2468 CTTAGTACCAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCTTTGGAGGATATAG 2527
 648 LysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsn 667
 2528 AAGCTAAATTTAACTCAATGCGCTTATTTGTGTGTGTATTTGTGCTATTAAT 2587
 668 GlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn 687
 2588 GGAGGCTCGCAACCCAGCATTTGCAATAATTTTCAAGATTTATAGGGTTTTTACAAGA 2647
 688 ---AsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIle 706
 2648 ATTGATGATCTCGAAACAAACAGCAGATAGTAACTGTTTCTACTATTTGTTCTAGCC 2707
 707 LeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGly 726
 2708 CTTGGAATTTATTTTATTTATTTTATTTTCTTTCAGGTTTTCACATTTGGCAAGCTGGA 2767
 727 GluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIle 746
 2768 GAGATCTCCACCAAGCGGCTCCGATACATGTTTCCGATCCATGCTCAGACAGGATGG 2827
 747 AlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAsp 766
 2828 AGTTGTTTCATGACCTTAACAAACACCATCTGGAGCATTGACTACCGGCTCGCCAATGAT 2887
 767 IleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThr 786
 2888 GCTGCTCAAGTTTAAAGGGCTATAGTTCAGGCTTGTGCTGTAATTAACCCAGATAATAGCA 2947
 787 AsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIle 806
 2948 AATCTTGGACAGGAATAATTATATCTTTCATCTATGTTGGTGGCAACTAACACTGTTTACTC 3007


```

Qy 1167 ArgThrClGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaasp 1186
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4088 AGAAGAGCGCGCACCTTGTCATTGTGATTGTCTACCGCCTGTCTCCACCATCCAGATGCAGAC 4147
Qy 1187 LeuIleValIleuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlnLeu 1206
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4148 TTAATAGTGTGTTCAGAAATGCGAGTCAAGGAGTCAGGAGCATGGCAGCATCAGAGCTGCTG 4207
Qy 1207 ArgAsnArgAspIleTyrPhelYsLeuValAsnAlaGlnSer 1220
::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4208 GCACAGAAGCATCTATTTTCAATGTCAGTGTCCAGGCT 4249

RESULT 9
US-09-120-513-1
; Sequence 1, Application US/09120513
; Patent No. 6025160
; GENERAL INFORMATION:
; APPLICANT: Brun, Kimberly
; APPLICANT: Chenery, Richard
; APPLICANT: Ellens, Harma
; APPLICANT: Field, John
; APPLICANT: Yue, Lin
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE
; TITLE OF INVENTION: SEQUENCES ENCODING RAT MDR1B2 AND
; TITLE OF INVENTION: SCREENING METHODS THEREOF
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY:
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: fastseq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/120.513
; FILING DATE: 22-JUL-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: King, William T
; REGISTRATION NUMBER: 30,954
; REFERENCE/DOCKET NUMBER: GPs0008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5015
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4233 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
US-09-120-513-1

Alignment Scores:
Pred. No.: 0 Length: 4233
Score: 3494.00 Matches: 684
Percent Similarity: 72.87% Conservative: 229
Best Local Similarity: 54.53% Mismatches: 274
Query Match: 56.73% Indels: 66
DB: 3 Gaps: 10

US-09-873-409-5 (1-1222) x US-09-120-513-1 (1-4233)

Qy 1 MetIleuGlyIleLeuIleAsnGlyLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 170 ATGGCTCTGGAACTCTCGCTGTGATTCACACGAACTGCTTCCTGCTGATGCTG 229

```



```
Db 2297 GTCATGGGAATGATTTCTTTTGTGTACGTACTCTTTCAAGGCTTCACATTTGGCAAGCT 2356
QY 726 GlyGluLeuLeuThrMetArgLeuArgHisLeuAlaPheLeuValAlaMetLeuTyrGlnAsp 745
Db 2357 GGAAGATCTCTACCAAGCAGCCCGATACATGTCTTCAATCCATGCTGCCAGCAGAT 2416
QY 746 IleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIle 765
Db 2417 ATAAGCTGGTTTGATGACCAATAAAACACCACTGGCTCGCTGACTACCAAGCTCGCTAGT 2476
QY 766 AspilealGlnIleGlnGlyValaThrGlySerArgIleGlyValLeuThrGlnAsnAla 785
Db 2477 GAGCGTTCTAATGTTAAAGGGCTATCGGCTCCAGGCTTCTGTAGTATCCCAAGATGTA 2536
QY 786 ThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeu 805
Db 2537 GCAAACTTGGCAGAGAAATATCTTATCTTGTAGTCTATGCTGGCAGCTTACACTTTTA 2596
QY 806 IleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaIleMetThr 825
Db 2597 CTGTGTAGTATATATACCACTCATTTGCTTGGTGGATATTGAAATGAAACTGTTGTCT 2656
QY 826 GlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAla 845
Db 2657 GGTCAAGCCTTGAAGGACAAAGAGAGATCTCTGGGAAGATCGCTACAGAAGCA 2716
QY 846 LeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyr 865
Db 2717 ATTGAAGAACTTCGCACTGTGTCTCTTGTAGTAGGAGCAGAGTTTGAACATATGTAT 2776
QY 866 GluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySer 885
Db 2777 GCCAGAGCTTGCAGATACCATACAGAAATGCTTTGAAGAAACACACAGCTTTGGATC 2836
QY 886 CysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGly 905
Db 2837 ACCTTCGCTTACCCAGGCGCATGATTTATTTTCTTATCTGCTGCTGTTTCCGCTCGGT 2896
QY 906 AlaTyrLeuIleGlnAlaGlyArgMetThrProGluClyMetPheIleValPheThrAla 925
Db 2897 GCCTACTTGTGGCAGAGAACTCATGACGTTTGAAGAAATGTTATGTTGGTATTTCTGCT 2956
QY 926 IleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLys 945
Db 2957 GTTGTCTTGTGTGTCATGCGCAGCAGGGAATACCATGTTTCATTCGCTCTGACTAGCGAAG 3016
QY 946 AlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuLysLysProAsnIleAspSer 965
Db 3017 GCCAAAGCTCTCAGCATCCCATCATCAGGATCATTCAGAGAAATCCCGAGATTGACAGC 3076
QY 966 ArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluVal 985
Db 3077 TACAGCAGGAGGCTTGAAGCCTTAATGTTGTAGAGAAATGTAATTAATGAGATC 3136
QY 986 SerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIle 1005
Db 3137 ATGTTCAACTATCCACCCGACCAACATCCAGTGTCTTCAGGAGCTGAGCTCGAGGTG 3196
QY 1006 GluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerVal 1025
Db 3197 AAGAAGGGGCAAGCTTCCGCTGCTGGGAGCAGTGGCTCGCGGAAGAGATACAGTGGTC 3256
QY 1026 GlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAsp 1045
Db 3257 CAGTGTCTGAGGCTTCTCAACCCCATGGCTGGAAACAGATGTTCTAGATGGCAAGAA 3316
QY 1046 AlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluPro 1065
Db 3317 ATAAACAACCTCAACGTCAGTGGCTCGCGCCACCTGGGCATGTTGTCCAGGAGCCC 3376
QY 1066 ValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValVal 1085
Db 3377 ATCCTGTTTGACTGCAGCATCACCGAGAACATCGCTTACGAGACAAACACCGCTGTCTGT 3436
```

```
QY 1086 ProLeuAspGluLeuLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGly 1105
Db 3437 TCTCATGAGGAGATCGTAGGGCCGCGCAGGAGGCCAACATCCACCATTCATCGACTCA 3496
QY 1106 LeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGln 1125
Db 3497 CTGCTCTGAGAAATACAAACACCAAGAGTGGGAGACAAAGGAGACTCAGCTGTGGGCGGCGAG 3556
QY 1126 LysGlnArgLeuAlaIleAlaArgAlaLeuGlnLysProLysIleLeuLeuAsp 1145
Db 3557 AAGCAGCGATGCCATCGCGCGCCCTCGTCAGACAGCCTCACATCTTACTTCTGGAT 3616
QY 1146 GluAlaThrSerAlaLeuAspAsnSerGluLysValValGlnHisAlaLeuAspLys 1165
Db 3617 GAAGCGACATCAGCTCTGGATCAGGAGAGTGAAAGAGTGTCCAGGAGCGCTGGACAAA 3676
QY 1166 AlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAla 1185
Db 3677 GCCAGGAAAGCGCACCTGCTGATTCAGAACGCCAGGTCAGAGACCGGCTGTCCACCATCCAGAACGCA 3736
QY 1186 AspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeu 1205
Db 3737 GACTTGTCTGTGATTCAGAACGCCAGGTCAGAGACCGGCTGTCCACCATCCAGAACGCA 3796
QY 1206 LeuArgAsnArgAspIleTyrPheLysLeuValAsnAla 1218
Db 3797 CTGCCCCAGAAAGCATCTATTCTCGATGGTTCAGGCT 3835
```

RESULT 10

```
US-09-450-105-1
; Sequence 1, Application US/09450105
; Patent No. 6169166
; GENERAL INFORMATION:
; APPLICANT: Kimberly Anne Brun
; APPLICANT: Richard James Chenery
; APPLICANT: Harma Ellens
; APPLICANT: John Anthony Feild
; APPLICANT: Lin Yue
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES
; TITLE OF INVENTION: ENCODING RAT MDR1B2 AND SCREENING METHODS THEREOF
; FILE REFERENCE: GP-5008-D1
; CURRENT APPLICATION NUMBER: US/09/450,105
; CURRENT FILING DATE: 1999-11-29
; EARLIER APPLICATION NUMBER: 09/120,513
; EARLIER FILING DATE: 1998-07-22
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 4233
; TYPE: DNA
; ORGANISM: HOMO SAPIENS
US-09-450-105-1
```

Alignment Scores:

Pred. No.:	0	Length:	4233
Score:	3494.00	Matches:	684
Percent Similarity:	72.87%	Conservative:	229
Best Local Similarity:	54.59%	Mismatches:	274
Query Match:	56.73%	Indels:	66
DB:	4	Gaps:	10

US-09-873-409-5 (1-1222) x US-09-450-105-1 (1-4233)

```
QY 1 MetIleGlyLysLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 170 ATGSGCTCTGGAACTCTCGCTGCTATCATCCAGAACCTGTCTCCCTCTCGTGTGCTG 229
QY 21 ValLeuGluMetSerAspAsnLeuIle----- 30
Db 230 GTTTCGGATACATGACAGATGTTTTTACCAGAGACAGCCGCAATTCGCGAGGCTT 289
QY 31 -----SerGlyCysLeuValGlnThrAsnThrTyrSerPheArgLeu----- 45
```


Db 290 ACTAATCAAGTGAATCAACAGTACACAGACCGTTCAGCGACAGCAGTCTGGAGGAGGAC 349
Qy 46 -----ThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db 350 ATGGCCATGTACGCTACTATTACACGGGCAATGGTGGCGGTGTCTCATCGTTCCTAC 409
Qy 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaAArgGlnThrIysArgIleArgIysGln 81
Db 410 ATCCAGGTTTCACTTTGGTGGCGGAGCTGGGAGACAAATACACAAAGATTAGGCAGAG 469
Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db 470 TTTTTCATGCCATCAATACAGAGATAGGCTGTGTTGACGTGAATGACGCTGGGAG 529
Qy 102 LeuAsnThrArgMetThr--AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db 530 CTCACACCCGGCTCACAGATGACGTCTCCAAATTAATGACGGAATTTGGTGCAAACTT 589
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db 590 GGAATGTTCTTTCAGTCCATAACGACATTTCAGCCGGTTTTATATAGGATTTATAAGT 649
Qy 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db 650 GGTGGAAGCTAACCTTGTAATTTGGCGCTCAGCCCTCTTATTGGGTGTCTCATCGC 709
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db 710 ATGTGGGCAAGGTACTGACTCTTATTACTTAATAAGAACTCCAGGCTTATGGAAGCT 769
Qy 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db 770 GGAGCAGTTGCCAAGAGCTTTAGCAGCCATCAGAACTGTGATTGGTGGAGGACAA 829
Qy 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
Db 830 AAGAAGCAACTTGAA----- 844
Qy 221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
Db 845 -----AGGTACAAATAAATAATTAGAGAA 868
Qy 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
Db 869 GCTAAAGAGTTGGCATAAAGAAAGCCATCACCGCAACATTTCCATAGTATGCTCTAC 928
Qy 261 PhePheMetAsnGlyTyrThrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
Db 929 CTGTTGCTATCGCTCTTATGCACCTGGCATCTGGTATGGGACCTCTTGGTCTCTCA 988
Qy 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
Db 989 AATGAA-----TATTCTATGGACAAGTCTTACCGTCTCTCTCTATTTATTGGGG 1042
Qy 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
Db 1043 ACTTTCAGTATTGGACATTTAGCCCCCAACATAGAGCCTTTGCAATGCAAGAGGGCA 1102
Qy 321 AlaPheHisIlePheGlnValIleAspLysPheProSerIleAspAsnPheSerThrAla 340
Db 1103 GCCTATGMAATCTTCAAGATAATTAATGATGAGCAAGCATTTGACAGCTTCTCAACCAAG 1162
Qy 341 GlyTyrLysProGluSerIleGlyGlyThrValGluPheLysAsnValSerPheAsnTyr 360
Db 1163 GGACACAAACACACAGATATAATGGGAAATTTGGAATTTTAAATAATTTACTTCAACTAC 1222
Qy 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
Db 1223 CCATCAGAAAGTGAAGTTAAGATCTTGAAGGGCTCAACCTGAAGGTGAAGAGCGGGCAG 1282
Qy 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400

Db 1283 ACGGTAGCCCTGTTGGCAACAGTGGCTGTGGGAAAGACAACTGTCTCAGCTGTGCGAG 1342
Qy 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
Db 1343 AGGCTCTACACCCCATAGAGGGCGAGGTTCAGTATCGACGGACAGGACATCAGGACCATC 1402
Qy 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
Db 1403 AATGTGAGGTATCTCGGGAAATCATTTGGGTGGTGGTTCAGTTCAGGAACCCGTGCTGTTGCC 1462
Qy 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
Db 1463 ACCACATTTGCCGAAACCAATTCGTATGGCCGCAAGAAACGTCCACATGATGATGATAGAG 1522
Qy 461 ArgAlaAArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
Db 1523 AAAGCTGTCAAGGAAGCCAATGCCATATGATCTCATCATGAACTGCCCCACAAATTAAC 1582
Qy 481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
Db 1583 ACCCTGTTGGTGAGAGAGGGCGCAGCTGAGTGGGGGACAGAAACAGAGGATCGCCATT 1642
Qy 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
Db 1643 GCCCGGCCCTGCTCGCAACCCCAAGATCCTTTTGTGTGATGAGCCACAGTCAGCCTTG 1702
Qy 521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThr 540
Db 1703 GACACAGAAGCGAAGCGCTGTTCAGGCGCTCTGGATAAGGCTAGAGAGGCGCGACC 1762
Qy 541 ThrIleValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560
Db 1763 ACCATTGTATAGTCCACCGCTTGTCTACAGTCGCAATGCTGACCTCATTTGCTGTTT 1822
Qy 561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580
Db 1823 GATGGTGGTGTCTATTGTGGAGCAAGGAATCATGAAGCTCATGAAGAGGAGGCAAT 1882
Qy 581 TyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMet 600
Db 1883 TACTTCAAACTTGTCTATGACACAG---ACTAGAGGAATGAAATTGACACAGGAATAAT 1939
Qy 601 ThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLys 620
Db 1940 GCTTATGAATCCCAAGTGACACTGGTGGCTCTGAGTTGACTTTCAGAAAAATCA---AAA 1996
Qy 621 SerAspPheIleAspLysAla-----GluGlu 629
Db 1997 TCTCTTTAATAGGAGATCAATTCGCAAGATATCCAGAGAGCAAGACACAGGAGAGA 2056
Qy 630 SerThrGlnSerLysGlu-----IleSerLeuProGluValSerLeuLysIle 646
Db 2057 AGACTTAGTTCGAAGAGATGTGGATGAAGATGTCCTATGCTTTCCTTTTGGCAGATC 2116
Qy 647 LeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeu 666
Db 2117 CTAAGCTAAATATTAGTGAATGGCCCTATTAGTTGTGGTGCTACTTTGTGCTGTATA 2176
Qy 667 AsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGly 686
Db 2177 AATGGTGTCAATCAACACAGTGTTCATAGTGTGTTTCAAGAGATTAGGGGTTTTTCA 2236
Qy 687 ---AsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheVal 705
Db 2237 AGAGACGACGACCATGAAACCAACCAACGAATGTAACTGTTTCCCTTCTCTTCTG 2296
Qy 706 IleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAla 725
Db 2297 GTCATGGGAATGATTTCTTTTGTAGTACTTCTTTCAAGGCTTTCATTTGCAAGACT 2356
Qy 726 GlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAsp 745
Db 2357 GGAGAGATCCTCACCAAGCGACTCCGATACATGCTCTTCAAAATCCATGCTGCGACAGAT 2416

QY 746 IleAlaTrpPheAspGluLysGluAsnSerThrGlyLeuThrThrIleLeuAlaIle 765
Db 2417 ATAAGCTGGTTGATGACCAATAAACCACACTGGCTCGCTGACTACCGAGCTCGTAGT 2476
QY 766 AspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAla 785
Db 2477 GAGCGCTTAAATGTTAAAGGGGCTATGGGCTCCAGGCTGCTAGTTTACCCAGAATGA 2536
QY 786 ThrAsnMetGlyLeuSerValIleIleSerPheIleTyrclyTrpGluMetThrPheLeu 805
Db 2537 GCAACCTTGGCAGAGAAATATCTTATCTTATCTGCTGCTGCGGAGCTTACACTTTTA 2596
QY 806 IleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThr 825
Db 2597 CTGTAGTATTATACCACTCATCTGCTGGGTGGAATATTGAATGAACCTGTTGTCT 2656
QY 826 GlyPheAlaAsnLysAspLysGlnLeuLysHisAlaGlyLysIleAlaThrGluAla 845
Db 2657 GGTCAAGCCTTGAAGGACAAAGAGCTAGAGATCTCTGGGAAGATCGCTACAGAAGCA 2716
QY 846 LeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTy 865
Db 2717 ATTGAATACTTCGCACTGTGTCTCTTCTGACTAGGAGCAGAAAGTTTGAACATATGTAT 2776
QY 866 GluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySer 885
Db 2777 GCCAGAGCTTGACATACCATACAGAAATGCTTGAAGAAAGCACACGCTTTGGGATC 2836
QY 886 CysTyAlaPheSerHisAlaPheIleTyPheAlaTyAlaAlaGlyPheArgPheGly 905
Db 2837 ACCTTCGCTTACCAGGCGCATGATTATTTTCTCTATGCTGCTGTTTCCGGTTCGGT 2896
QY 906 AlaTyIleuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAla 925
Db 2897 GCCTACTGTGGCCAGCAGAACTCATGACCTTGAATAATGTTATGTTGTTATTTCTGCT 2956
QY 926 IleAlaTyGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTySerLys 945
Db 2957 GTTGTCTTGGTGCCATGGCAGCAGGAATACCATGTCCTCTGACTAGCGGAAG 3016
QY 946 AlaLysSerGlyAlaAlaHisLeuPheAlaLeuGluLysLysProAsnIleAspSer 965
Db 3017 GCCAAAGCTCAGCATCCCATCATCAGCATCATGAGAAATCCCGAGATTGACAGC 3076
QY 966 ArgSerGlnGluGlyLysProAspThrCysGluGlyAsnLeuGluPheArgGluVal 985
Db 3077 TACAGCACGGAGGCTTGAAGCCCTTAATTTGTTAGAGGAAATGTGAATTTAATGAGTC 3136
QY 986 SerPhePheTyProCysArgProAspValPheIleLeuArgGlyLeuSerIle 1005
Db 3137 ATGTTCAACTATCCACCACCAACATCCAGTGTCTCAGGAGCTGAGCTTCGAGGTG 3196
QY 1006 GluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerVal 1025
Db 3197 AAGAAGGGCAAAACGCTTCGCTGTGGCAGCAGTGGCTCGCGGAAGAGTACAGTGGTC 3256
QY 1026 GlnLeuGlnArgLeuTyArgProValGlnGlnValLeuPheAspGlyValAsp 1045
Db 3257 CAGCTGCTCAGCGCTTCTCAACCCCATCGCTGGAACAGTGTCTTAGATGGCAAGAA 3316
QY 1046 AlaLysGluLeuAsnValGlnTTPLeuArgSerGlnIleAlaIleValProGlnGluPro 1065
Db 3317 ATAACAACCTCAACGTCAGTGGCTCGCGCCACCTCGGCGATTTGTCCAGGAGCCC 3376
QY 1066 ValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyArgGlyAspAsnSerArgVal 1085
Db 3377 ATCCTGTTTGCATCGCAGCATCCCGGAGCAACATCGCTTACGAGAGCAACAGCCGTGCTG 3436
QY 1086 ProLeuAspGluIleLysGluAlaAlaAsnAlaAsnIleHisSerPheIleGluGly 1105
Db 3437 TCTCATGAGGATCGTGAGGGCGCGCCAGGAGGCCAACATCCACCAAGTTTCATCGACTCA 3496

QY 1106 LeuProGluLysTyAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGln 1125
Db 3497 CTGCTGTAGAAATACAAACACAGAGTGGGAGACAAAGGACTCAGCTGTGCGGCGGCAG 3556
QY 1126 LysGlnArgLeuAlaIleAlaArgAlaLeuGlnLysProLysIleLeuLeuAsp 1145
Db 3557 AAGCAGCGATCGCCATCGCGCGCCCTCGTCAGACAGCCCTACATCTTACTTCTGGAT 3616
QY 1146 GluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLys 1165
Db 3617 GAAGCGACATCAGCTCTGATACGGAGATGAAAGGTGCTCCAGGAAGCGCTGGACAA 3676
QY 1166 AlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAla 1185
Db 3677 GCCAGGAAGCGCGACCTGCTGATCGCGCACCGGCTGTCCACCATCCAGAAGCGCA 3736
QY 1186 AspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeu 1205
Db 3737 GACTTGATCGTGTGATTGAGAACCGGCCAGGTCAAGGAGCAGCGGCCACCCAGCAGCTG 3796
QY 1206 LeuArgAsnArgAspIleTyPheLysLeuValAsnAla 1218
Db 3797 CTGCCCCAGAAAGCATCTATTCTCGATGTTTCAGGCT 3835
RESULT 11
US-08-583-276-18
; Sequence 18, Application US/08583276
; Patent No. 5837536
; GENERAL INFORMATION:
; APPLICANT: McDonagh, Kevin T.
; APPLICANT: Nienhuis, Arthur
; APPLICANT: Tolstoshev, Paul
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HUMAN
; TITLE OF INVENTION: MULTIDRUG RESISTANCE GENES AND IMPROVED
; TITLE OF INVENTION: SELECTION OF CELLS TRANSDUCED WITH SUCH GENES
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan,
; ADDRESSEE: Cecchi & Stewart
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: DW4.V2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/583,276
; FILING DATE: 05-JAN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/332,444
; FILING DATE: 31-OCT-1994
; APPLICATION NUMBER: 07/887,712
; FILING DATE: 22-MAY-1992
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4669 bases
; TYPE: nucleic acid
; STRANDEDNESS: singular
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: Genomic DNA
US-08-583-276-18
Alignment Scores:
Pred. No.: 0 Length: 4669
Score: 3473.50 Matches: 673
Percent Similarity: 73.34% Conservative: 246
Best Local Similarity: 53.71% Mismatches: 271

Query Match: 56.40% Indels: 63
DB: 2 Gaps: 10
US-09-873-409-5 (1-1222) x US-08-583-276-18 (1-4669)

QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyValaCysLeuProLeuMetProLeu 20
DB 575 ATGGTGGTGGAACTTGGCTGCATCATCCATGGGGCTGGACTTCTCTCATGATGCTG 634
QY 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
DB 635 GTGTTTGGAGAAATGACAGATATCTTTGCAAAATGCAGGAATTTAGAGAGATCTGATGCTCA 694
QY 32 GlyCysLeuValGlnThr-----AsnThrTySerPhePhe----- 43
DB 695 AACATCACTAATAAGAGTATATCAATGATACAGGGTCTTCTCAATCTGGAGGAAGAC 754
QY 44 -----ArgLeuThrLeuTyTyValGlyIleGlyValaAlaLeuIlePheGlyTy 61
DB 755 ATGACCAAGGTATGCCATTATTAACAGTGAATTTGGTGTGGGTGCTGGTGTCTGCTTAC 814
QY 62 IleGlnIleSerLeuTrpIleThrAlaAlaArgGlnThrLysArgIleAqLysGln 81
DB 815 ATTCAAGTTTCATTGGTCCCTGGCAGCTGGAGCAATACACAAATATAGAAACAG 874
QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
DB 875 TTTTTCATGCTATAATCGCACAGGAGATAGGCTGTTGATGTCACGATGTTGGCAG 934
QY 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
DB 935 CTTAAACCCCGCTTACAGATGATGCTCTTAAGATTAAATGAATTTATGGTGACAAAT 994
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
DB 995 GGAATGTTCTTCAGTCAATGGCAACATTTTCTACCTGGGTTTATAGTAGATTACACGT 1054
QY 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
DB 1055 GGTGGGAAGCTAACCTTGTGATTTTGGCCATCAGTCTCTTCTGGCAGTGTGCTGCT 1114
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTySerLysAla 180
DB 1115 GTCTGGGCAAGATACTATCTTCTTACTATAAGAACTCTTAGCGTATGCAAAAGCT 1174
QY 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
DB 1175 GGACGAGTAGCTGAAGAGGCTTGGCAGCAATTAGAACTGTGATGATTTGGAGGACAA 1234
QY 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyAlaTrpPheTyPhe 220
DB 1235 AAGAAAGAACTTGAA----- 1249
QY 221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyThrGlnAsnLeuLysAsp 240
DB 1250 -----AGGTACACACAAATAATTTAGAGAA 1273
QY 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTy 260
DB 1274 GCTAAAGAAATTTGGGATAAAGAGCTTATACAGCCCAATATTTCTATAGGTGCTGCTTC 1333
QY 261 PhePheMetAsnGlyTyThrTyGlyLeuAlaPheTrpTyGlyThrSerLeuIleLeuAsn 280
DB 1334 CTGCTGATCTATCATCTTATGCTCTGGCTTCTGGTATGGGACCCTTGGTCTCTCA 1393
QY 281 GlyGluProGlyTyThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
DB 1394 GGGGAA-----TATTCTATTGGCAAGTACTCACTGATTTCTTTCTGTTAATTTGGG 1447
QY 301 SerTyrcysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
DB 1448 GCTTTTAGTGTGGACAGGCATCTCCAAGCATTTGAAGCATTTTGCAAATGCAAGAGGACGA 1507

QY 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
DB 1508 GCTTATGAATCTTCAAGATAATTTGATAATAAGCCCAAGTATTGACAGCTATTGCAAGAGT 1567
QY 341 GlyTyLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTy 360
DB 1568 GGGCACAAACAGATATAATTAAGGGAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 1627
QY 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
DB 1628 CCATCTCGAAAAAGACTTAAGATCTTGAAGGGCTGAACTGAAGCTGCAGAGTGGCAG 1687
QY 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValGlnLeuLeuGln 400
DB 1688 ACGTCCCCCTGGTTGGAAACAGTGGCTGTGGAAAGACACACACAGTCCAGCTGATGTCAG 1747
QY 401 ArgLeuTyAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
DB 1748 AGGCTCTATGACCCACAGAGGGGATGGTCAGTGTTCATGGACAGGATATTAGGACCATA 1807
QY 421 AsnValArgHisTyArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
DB 1808 AATGTAAGGTTTCTACGGGAATCAATGGTGTGGTCAGTCAGGAACCTGTATTGTTGCC 1867
QY 441 ThrThrIleSerAsnAsnIleLysTyArgAspAspValThrAspGluGluMetGlu 460
DB 1868 ACCACGATAGCTGAAACATTCGCTATGGCCGTGAAATGTCCACCATGATGATGATGAG 1927
QY 461 ArgAlaAlaArgGluAlaAsnAlaTyAspPheIleMetGluPheProLeuLysPheAsn 480
DB 1928 AAGCTGTCCAGAGACCAATGCCATGATCTTATCATGAACCTGCTCTAATAATTGAC 1987
QY 481 ThrLeuValGlyLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
DB 1988 ACCCTGTTGGAGAGAGGGCCCGCTGAGTGTGGTGGCAGAGCAGAGGATCGCCATT 2047
QY 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
DB 2048 GCACGTGCCCTGTTGCGAACCCCAAGATCTCTCTGCTGGATGAGGCCAGCTGAGCCTTG 2107
QY 521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThr 540
DB 2108 GACACAGAAGCCAGACAGTGGTTCAGTGGCTCTCGATAAGGCCAGAAAGTCCGACC 2167
QY 541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560
DB 2168 ACCATTGTGATGCTCATCGTTGTCTACAGTTCGTAACTGTCACGCTCGCTGCTTTC 2227
QY 561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580
DB 2228 GATGATGGAGTCAATTGTGGAGAAAGGAATCATGATGAACCTCATGAAGAGAAAGCAT 2287
QY 581 TyrTySerLeuValMetSerGln-----AspIleLysLysAlaAsp 594
DB 2288 TACTTCAAACTTGTCAATGACAGACAGCAGGAAATGAAGTTGAATTTAGAAAATGCGAGCT 2347
QY 595 GluGlnMetGlnSerMetThrTySerThrGluArgLysThrAsnSerLeuProLeuHis 614
DB 2348 GATGAATCAAAAGTGAATTTGATGCTTGGAAATGCTCTCAATGATTCAGATCACGT 2407
QY 615 SerVal-----LysSerIleLys---SerAspPheIleAspLysAla 627
DB 2408 CTAATAAGAAAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAACCCCAAGACAGAAAG 2467
QY 628 GluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeu 647
DB 2468 CTTAGTACCAAAAGAGCTCTGGATGAAGATATACCTCCAGCTTCTTTTGGAGGATATG 2527
QY 648 LysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsn 667
DB 2528 AAGCTAAATTTAACTGAATGGCTTATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2587
QY 668 GlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn 687

REFERENCE/DOCKET NUMBER: PA-0012 CIP 2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301 527-2058
TELEFAX: 301 208-6997
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2726 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-461-823-1

Alignment Scores:

Pred. No.: 1,31e-245 Length: 2726
Score: 2280.00 Matches: 435
Percent Similarity: 76.55% Conservative: 159
Best Local Similarity: 56.06% Mismatches: 168
Query Match: 37.02% Indels: 14
DB: 1 Gaps: 4

US-09-873-409-5 (1-1222) x US-08-461-823-1 (1-2726)

QY 459 MetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLys 478
DB 2 ATTGAGAAAGCTGTCAAGGAGCGCAATGCTATGACTTTATCATGAACCTGCCTCATAAA 61
QY 479 PheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIle 498
DB 62 TTTGACACCTGTTGGAGAGAGAGGGGCCAGTTCAGTGGTGGCAGAGCAGAGGATC 121
QY 499 AlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSer 518
DB 122 GCATTTGCACGTGCTGCTGCAACCCCAAGATCTCTGCTGGATGAGGCCACGTCA 181
QY 519 AlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGly 538
DB 182 GCCTTGACACAGAAAGCGAAGCAGTGGTTCAGGTGGCTCGGATAAGCCAGAAAGGT 241
QY 539 ArgThrThrIleValAlaAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleVal 558
DB 242 CGGACCAACCATTTGTGATGCTCATCGTTTGTCTACAGTTCGTAAATGCTGACGTCACTCG 301
QY 559 ThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArg 578
DB 302 GGTTCGATGATGGATGCTATTGGGAGAAAGGAAATCATGATGAACCTCATGAAAGAGAAA 361
QY 579 GlyLeuTyrTyrSerLeuValMetSerGln-----AspIleLysLys 592
DB 362 GGCATTTCCTCAAACTTGTCACAATGCAGACAGCAGGAAATGAAGTTGAATTAGAAAT 421
QY 593 AlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuPro 612
DB 422 GCAGCTGATGAATCCAAAGTGAATGATGCTTGGAAATGCTTCAAAATGATTCAAGA 481
QY 613 LeuHisSerVal-----LysSerIleLys---SerAspPheIleAsp 625
DB 482 TCCAGTCTTAATAGAAAAAGATCAACTCGTAGGAGTGTCCGTGGATCAACAAGCCCAAGAC 541
QY 626 LysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLysLys 645
DB 542 AGAAACCTTAGTACCAAGAGGCTCTGGATGAAGATATACCTCCAGTTTCCTTTGGAGG 601
QY 646 IleLeuLysLeuAsnLysProGluTyrProPheValValLeuGlyThrLeuAlaSerVal 665
DB 602 ATTATCAAGCTAAATTTAACTGAATGGCTTATTTTGTGTTGTTGTTGTTGTTGTTGTT 661
QY 666 LeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPhe 695
DB 662 ATAAATGGAGGCTGCAACACGACATTTGCAATATATTTTCAAAAGATTATAGGGGTTTT 721

QY 686 GlyAsn---AsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePhe 704
DB 722 ACAAGAAATTTGATGATCTCTGAAACAAAACGACAGAAATAGTAACCTTGTTTTCACTATTGTTT 781
QY 705 ValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArg 724
DB 782 CTAGCCCTTGAATATTCTTTTATTATCATTTTCTTCAGGGTTTCACATTGGCAAA 841
QY 725 AlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGln 744
DB 842 GCTGGAGAGATCCTCACCAAGCGGCTCCGATACATGGTTTCCGATCCATGCTCAGACAG 901
QY 745 AspIleAlaTyrPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAla 764
DB 902 GATGTGAGTGGTTTGTATGATGACCTTAAAAACACCACTGGGAGCATTGACTACCGGCTCGCC 961
QY 765 IleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsn 784
DB 962 AATGATGCTGCTCAAGTTAAAGGGCTATAGGTTCAGGCTTCTGTAATATCCCAAGAT 1021
QY 785 AlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyrGluMetThrPhe 804
DB 1022 ATAGCAATCTTGGGACAGGAATATATATCTTCATCTATGTTGGCACTAACACTG 1081
QY 805 LeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMet 824
DB 1082 TTAATCTTGAATTTGATACCCATCTTCAATAGCAGGAGTGTGTTGAAATGAAATGTTG 1141
QY 825 ThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGlu 844
DB 1142 TCTGGCAAGCACTGAAGATGAAGAAAGAACTAGAAAGTGTGGGAGGATCGCTACTGAA 1201
QY 845 AlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMet 864
DB 1202 GCATAGAAAACTTCGAAACCGTGTGTTCTTTGACTCAGACAGCAAGTTGAAACATATG 1261
QY 865 TyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleGly 884
DB 1262 TATGCTCAGAGTTTGCAGGTACCATACAGAAACTCTTTCAGGAAAGCACACATCTTTGGA 1321
QY 885 SerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPhe 904
DB 1322 ATTACATTTTCTTCCACCGCAATGATGATATTTTCTATGCTGGATGTTTCCGGTTT 1381
QY 905 GlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGluMetPheIleValPheThr 924
DB 1382 GGAGCTACTTGGTGGCACAATAAATCATGAGCTTTGAGGATGTTCTGTAGTATTTTCA 1441
QY 925 AlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSer 944
DB 1442 GCTGTTGTTGTTGGTGGCATGCCGCGGGGCAAGTCAGTTCTCTGCTGACTATGCC 1501
QY 945 LysAlaLysSerGlyAlaAlaHisIlePheAlaLeuLeuGluLysLysProAsnIleAsp 964
DB 1502 AAAGCCAAATATCAGCAGCCACATCATCATGATCATTTGAAATAAAACCCCTTTGATTGAC 1561
QY 965 SerArgSerGluGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGlu 984
DB 1562 AGCTACAGCAGGAGGCTTAATGCCGAACACATTTGGAAGGAAATGTACATTGTTGTTGAA 1621
QY 985 ValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSer 1004
DB 1622 GTTGTATTCAACTATCCCAACCGGACATCCAGTCTTTCAGGACTGAGCCTGGAG 1681
QY 1005 IleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSer 1024
DB 1682 GTGAAGAGGGCCAGACGCTGCTGTGTGGCAGCAGTGGCTGTGGGAGACACAGTG 1741
QY 1025 ValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlnValLeuPheAspGlyVal 1044
DB 1742 GTCCAGCTCTCGAGCGGTTCTACGACCCCTTGGCAGGAAAGTGTGCTTGTGTCAAA 1801
QY 1045 AspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGlu 1064


```
Db 3202 CGAAGCCACAAATTGATACTGGTCTGAAGAGGGCGAGAGGCTCGAAGCGTGGAGGTT 3261
Qy 979 AsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeu 998
Db 3262 GAATCGAATTTAGAACGTGCCTTCAGATACCGACCGCCGACAGACAGCTGCTG 3321
Qy 999 ArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGly 1018
Db 3322 CGCGCTTGACCTGACCGTGAACCTGGACATATGTCGGCTTGTCGACCCAGCGGT 3381
Qy 1019 CysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGln 1038
Db 3382 TGTGGCAAGATGACCACTGCTGCTGAGCGCTTTTACGATGCGATGCGGCTCC 3441
Qy 1039 ValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIle 1058
Db 3442 ATCTTGTGATGGAGGACATAGTAACTAAATATCACTTCACTACCGACGCTTCTG 3501
Qy 1059 AlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyr 1078
Db 3502 TCATGCTGACCGACGAGCCGACACTGTACCGGCGCCATCAAGGAAACATCTTACTT 3561
Qy 1079 GlyAspAsnSerArgValValProLeuAspGluLeuLysGluAlaAlaAsnAlaAsn 1098
Db 3562 GGTATTGCGAAGATGACGTACCGGAAGAAATCTTGATTAGGCTTGCAAGGACGCTAAT 3621
Qy 1099 IleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGly 1118
Db 3622 ATTACGACTTCATCATGCTGCTCCCGAGGGCTTTAATACAGTTGTGCGACGAGGA 3681
Qy 1119 AlaGlnLeuSerGlyGlnGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLys 1138
Db 3682 GGCATGTTGCTGCGCGCCAAAGACGAGTGTCCTATTCGCCGAGCCCTTCTCGGAT 3741
Qy 1139 ProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysVal 1158
Db 3742 CCCAAATCTCTTCTCGATGAGCGACGCTCAGCCCTCGACTCCGAGTCAGAAAAGGTC 3801
Qy 1159 ValGlnHisAlaLeuAspLysAlaAArgThrGlyArgThrCysLeuValValThrHisArg 1178
Db 3802 GTCCAGCGCGCTTGGATGCGCTGCGCGAGCGCCGAACCAATCGCGCTTGACACCGA 3861
Qy 1179 LeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys 1198
Db 3862 CTCAGCACGATTCBAAGCGCGAGTTATCTATGTTTTCGACCAAGCGACGTCGAA 3921
Qy 1199 GlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAla 1218
Db 3922 AGCGGAACGACACGCACTGCTCCAGAAAAGGCGCGTACTACGAGCTGGTCAACTG 3981
Qy 1219 GlnSerVal 1221
Db 3982 CAGAGCTTG 3990
```

RESULT 14

```
US-08-996-545-3
; Sequence 3 Application US/08996545
; Patent No. 5928898
; GENERAL INFORMATION:
; APPLICANT: Skatrud, Paul L.
; APPLICANT: de Waard, Maarten A.
; APPLICANT: Peery, Robert B.
; APPLICANT: Andrade, Alan C.
; TITLE OF INVENTION: Multiple Drug Resistance Gene atRD of
; TITLE OF INVENTION: Aspergillus nidulans
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: U.S.
; ZIP: 46285
```

```
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/996,545
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Webster, Thomas D.
REGISTRATION NUMBER: 39,872
REFERENCE/DOCKET NUMBER: X-11766
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-3334
TELEFAX: 317-276-2763
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 4002 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-996-545-3

Alignment Scores:
Pred. No.: 6,01e-243 Length: 4002
Score: 2259.00 Matches: 494
Percent Similarity: 58.41% Conservative: 232
Best Local Similarity: 39.74% Mismatches: 446
Query Match: 36.68% Indels: 72
DB: 2 Gaps: 12

US-09-873-409-5 (1-1222) x US-08-996-545-3 (1-4002)
```

```
Qy 15 LeuProLeuMetProLeuValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu 34
Db 368 GUGCCAUUCUGCGCGGUCGCUUCCAG------GGAUAUUGUUAU 409
Qy 35 ValGlnThrAsnThrTyrSerPhePhe-----ArgLeuThrLeuTyrVal 50
Db 410 AUCAAAUUCUGUACGAC-GAGUUCUAUGAUAUACCAAGAACGACGUAUCUUAUCC 468
Qy 51 GlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeuTrpIleThr 70
Db 469 UACUCUGUAUCGCGAGUUUGUCACUGCUAUGUUAUACUGUUGGCUUUAUUAUCC 528
Qy 71 AlaAlaArgGlnThrLysArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAsp 90
Db 529 GGAGAACACGCCACGACGAGAACUCCGCGUAUUAUACUUGUCUUAUCCGCGCAGAAC 588
Qy 91 IleGlyTrpPheAspSerCysAspIleGlyGluLeuAsnThrArgMetThr---AspIle 109
Db 589 AUUGGCUAUUUUGAUAACUCUGGUGCGGGAAGUACACCGCUAUAACAGCGCAUACA 648
Qy 110 AspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGlnAsnMetSerThr 129
Db 649 AACCUUAUCCAGGAUGGCAUUCGGAAGGUGGUGGUCUCUUUGACUGCCCGGCGACA 708
Qy 130 PheSerIleGlyLeuAlaValGlyLeuValLysGlyTrpLysLeuThrLeuValThrLeu 149
Db 709 UUGUGACAGCAUUAUUAUCCGCUAUAUACGUAUUAUUAUUAUUAUUAUUAUUAUUA 768
Qy 150 SerThrSerProLeuIleMetAlaSerAlaAlaAlaCysSerArgMetValIleSerLeu 169
Db 769 UCAACAUAUUGGCGCCUCUGUCUACCAUGGGCGGUGGUGUCUCAGUUAUUAUUAUUA 828
Qy 170 ThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluValLeuSer 189
Db 829 AGCAAAAAGUUGGCUUAGCAGCUACGUGGCGGCGACUGUUGCGGAGAGGUCUACGAC 888
```


QY 190 SerIleArgThrValIleAlaPheArgAlaGlnGlnLysGlnLeuGlnArgSerPheLeu 209
DB 889 UCCAUCAAGAAUCCGACAGCGGUUGGACCCACCAAGACAAGCUGCGAAG----- 936
QY 210 LeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCysValLeu 229
DB 936 ----- 936
QY 230 ***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThr 249
DB 937 -----CAGUAGAGGUCCACUUGAGCAAGAGCUGAGAAUUGGGGAACAAAGAACCCAG 987
QY 250 IleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeu 269
DB 988 AUUGUCAUGGGUUCUAGAUGGCGCAUGUUGGGCCUUAUAGUACUAGCAACUAGCGGUU 1047
QY 270 AlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThr 289
DB 1048 GGCUCUGGAUGGUUCUGUUUCCUGGUGAGUGGU-----GCAGUCCAUUGGGUGAU 1101
QY 290 ValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValPro 309
DB 1102 AUUCUCACAGUUCUAGCGCAUCUUGAGUAGCGGUAUCUUCUUGGGGAACGUUAGUCCA 1161
QY 310 HisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAsp 329
DB 1162 AAUGCUCAGCAUUUACAACGCGUGGGCGCGCGCAAGAAUUAUUGGAACGAGUCCAU 1221
QY 330 LysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGly 349
DB 1222 CGCAGUCCCAUUAUAGUCCAUUAUUGCAAGAGGGAAGACGCGUCCAGCAUUUUGAGGGC 1281
QY 350 ThrValGlnPheLysAsnValSerPheAsnTyrProSerArgProSerIleLysLeu 369
DB 1282 CACAUUGAGUUAUCCGAUAGUACGAUUAUUAUCCCAUAGACCGGAGGUCACCGUACUG 1341
QY 370 LysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGly 389
DB 1342 GAGGAUGUUUCUGUCAACGCGCGCGGAAAAACAACCGUUUAGTGCGCCCCUCUGGC 1401
QY 390 SerGlyLysSerThrValValGlnLeuGlnArgLeuTyrAspProAspAspGlyPhe 409
DB 1402 UCUGGAAAAAGUACGGUGGCGCGUUGGUGGAGCAUUCUACAGUCCUGUCCGUGUACG 1461
QY 410 IleMetValAspGlnAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIle 429
DB 1462 GUUUUCUGAGUCCGCAUGACAUCAGGACCUCAUUCUGCGUGGUCGCCAACAGAU 1521
QY 430 GlyValIleSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyr 449
DB 1522 UCUUUGUUGAGCCAGGAGCGGUUCUUUUGGCACGACGAUUAUAAAGAAUUAUAGGCAC 1581
QY 450 Gly-----ArgAspAspValThrAspGlnGluMetGlu 460
DB 1582 GGUCUCAUGCGCACAAAGUACGAGAAUAGAAUCCGAGGAUAGGUCGCGGAACUACUGCAG 1641
QY 461 ArgAlaAlaArgGlnAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
DB 1642 AACCGCGCAAAAUAGCGGAUUGCUAGCUAGCUUUAUUAUCCUUGCCUGGAGGUUAUGAG 1701
QY 481 ThrLeuValGlyLysGlyAlaGlnMetSerGlyGlnLysGlnArgIleAlaIle 500
DB 1702 ACCAAUUGGGCAGCGGCGUUUCUUUCCUACGAGGUGGCCAGAAACAGCGCAUUGCAU 1761
QY 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGlnAlaThrSerAlaLeu 520
DB 1762 GCCCGGCGGUUGUAGUAGACCCAAAAAUUCUGUCCUGGAUAGCAUUCUGCGGU 1821
QY 521 AspSerGluSerLysSerAlaValGlnAlaLeuGluLysAlaSerLysGlyArgThr 540
DB 1822 GACAAAAUCCGAGCGGUGGUUAGAGCGCUUUGAGGGCGGCAAGCCGAACU 1881
QY 541 ThrIleValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560

DB 1882 ACUAUUGUAGUCCGUAUCGCUUCCAGCAUCAAACCGCCACCAACAUUGUGGUUUG 1941
QY 561 LysAspGlyMetLeuAlaGlnLysGlyAlaHisAlaGlnLeuMetAlaLysArgGlyLeu 580
DB 1942 GUCAUUGGCAAAAUUGUGUACAAGGAACUCACGAUAGUUGUUGUACCGCGGAGCGCU 2001
QY 581 TyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGlnMetGluSerMet 600
DB 2002 UAUCGCAAAUUGUGGAGGCUACGUAUCAAUAGAACAAGGAAGCUGACGCGCUUGAG 2061
QY 601 ThrTyrSerThrGluArgLysThrAsnSer-----LeuProLeuHisSerValLysSer 618
DB 2062 GACGCGACGUGAGUACUCACGAUAGCAGAUUUGCCAAAUACAACUGCGUCAAAC 2121
QY 619 IleLysSerAspPhe-----IleAspLys-----AlaGluGluSer 630
DB 2122 GCANAUCCGACUUCGACGGAACAAACCAACCAUUGACCGGACCGGCCACCAAGUCU 2181
QY 631 ThrGlnSerLysGluIleSer-----LeuProGluValSerLeuLeu 644
DB 2182 GUUUCAGCGCGAUUCUUUAAAAGACCCCGGAAACAACUCCGAAAUACUACUUAUGG 2241
QY 645 LysIleLeuLys-----LeuAsnLysProGluTrpProPheValValLeuGly 660
DB 2242 ACGCUGCUCAAAUUGUUGUCCUUAACCGCCCGGAAUCCCGUACUUGCUACUGCU 2301
QY 661 ThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLys 680
DB 2302 CUUGUCUUCUAGUUGUUGGUGGUGGCGCAACCCCAAGCAGGUGCUUAUUGCUAAA 2361
QY 681 IleIleThrMetPheGly-----AsnAsnAspLysThrThrLeuLysHisAspAlaGlu 698
DB 2362 GCCAUCAGCACACUCGCGUCCAGAAUACAAUUAUAGCAAGCUUCGACAUAGCGGAU 2421
QY 699 IleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGln 718
DB 2422 UUCUGUCAUUGAUUGUUCUGUGGUGUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 2481
QY 719 GlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPhe 738
DB 2482 GUGUGCAUUGCCGUAUUGUCCGAGAGACUUAUUGCGCGCGAGAGACACUGCUUU 2541
QY 739 LysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGlnLysGlnAsnSerThrGlyGly 758
DB 2542 CGGACGAUACUCCGUAACAGAUUGCUUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 2601
QY 759 LeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIle 778
DB 2602 CUGACCUUUCUGUCCACCGAGACGAAGCAUCUCUGCGUGUUAUUAUUAUUAUUAUUA 2661
QY 779 GlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyr 798
DB 2662 GGCAUGAUUGAGUAGACCUCCAGACCGCGGAGGCGGCUUAUUAUUAUUAUUAUUAUUA 2721
QY 799 GlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMet 818
DB 2722 GGGUGAAAUUGCCUUAUUGUUAUCUGUGUUGCGGUGUCCUGGCAUGCGGUUUC 2781
QY 819 IleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAla 838
DB 2782 UAACGAUUAUUAUUGUAGCCCAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 2841
QY 839 GlyLysIleAlaThrGlnAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGlu 858
DB 2842 GCAACAUUGUGGAGGCUACUUGCUUAUCCGACAGUUGCGUUAUUAUUAUUAUUAUUA 2901
QY 859 LysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLys 878
DB 2902 AGGGAUUGCGGGAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 2961
QY 879 LysAlaGlnIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyr 898

Db 529 GGAGAACACGCCACGAGATCCGCGAGTATTACTTGAGTCTATCTCGCGCCAGAAC 588
Qy 91 IleGlyTrpPheAspSerCysAspIleGlyGluLeuAsnThrArgMetThr---AspIle 109
Db 589 ATTGGCTATTTTGATAACTCGGTGCGGGAGTACACCCCGTATACACGCCGATACA 648
Qy 110 AspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGlnAsnMetSerThr 129
Db 649 AACCTTATCCAGGATGGCATTTGCGAGAGGTGCGTCTCACTTTGATCGCCCTGGCGACA 708
Qy 130 PheSerIleGlyLeuAlaValGlyLeuValLysGlyTrpLysLeuThrLeuValThrLeu 149
Db 709 TTGTCGACAGCATTCATTATCGGCTACGTAATATCTGGAAGTTGGCTCTAATTTCCAGC 768
Qy 150 SerThrSerProLeuIleMetAlaSerAlaAlaAlaCysSerArgMetValIleSerLeu 169
Db 769 TCAACAATTGTGSCCCTCGTTCTCACCATGGGGGGTGTCTCAGTTTATCATCAAGTAC 828
Qy 170 ThrSerLysGluLeuSerAlaTrpSerLysAlaGlyAlaValAlaGluGluValLeuSer 189
Db 829 AGCAAAAGTCGCTTGACAGCTACCGTGCAGCGGCACCTGTTGCGAAGAGGTTCATCAGC 888
Qy 190 SerIleArgThrValIleAlaPheArgAlaGlnGluLysLeuGlnArgSerPheLeu 209
Db 889 TCCATCAGAAATGCCACAGCGTITGGCCCAAGCAAGCTTCGGAAG----- 936
Qy 210 LeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCysValLeu 229
Db 936 ----- 936
Qy 230 ***PheValargTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThr 249
Db 937 -----CAGTATGAGGTGCCACTTAGACGAAGCTGAGAAATGGGGAACAAAGAACCCAG 987
Qy 250 IleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeu 269
Db 988 ATTGTCATGGTTTCATGATGGCGCCATGTTTGGCCCTTATGACTCGAACTACGGTCTT 1047
Qy 270 AlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThr 289
Db 1048 GGCTTCTGATGGTGTCTCGTTTCTGGTAGATGGT-----CGAGTCGATGGGTGAT 1101
Qy 290 ValLeuAlaValPhePheSerValIleHisSerTyrCysIleGlyAlaAlaValPro 309
Db 1102 ATTCTCACAGTCTCATGGCCATCTTGATCGGATCGTTCTCTTGGGAACGTTAGTCCA 1161
Qy 310 HisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAsp 329
Db 1162 AATGCTCAAGCATTTACAACGCTGTGGCGCGCGCAAGATATTTGGAACGATCGAT 1221
Qy 330 LysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGly 349
Db 1222 CGCAGTCCCATAGATCCATATTCGACGAAGGGAAGACGCTCGACCNATTTTGAGGGC 1281
Qy 350 ThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeu 369
Db 1282 CACATTGAGTTACGCAATGTCAAGCATATTATCCCATCTAGACCCGAGGTCAACGTCATG 1341
Qy 370 LysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGly 389
Db 1342 GAGGATGTTTCTCTGTCATCGCCGCTGGAAAAACAACCGCTTTAGTCGCCCTCTGGC 1401
Qy 390 SerGlyLysSerThrValValGlnLeuGlnArgLeuTyrAspProAspAspGlyPhe 409
Db 1402 TCTGGAAGAAAGTACGGTGGTGGCTTGGTTGAGCGATTTACATGCTGTTCGCGGTACG 1461
Qy 410 IleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIle 429
Db 1462 GTTTTGTGCGATGGCCATGACATCAAGGACCTCAATCTCGCTGGCTTCGCCAACAGATC 1521
Qy 430 GlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnIleLysTyr 449
Db 1522 TCTTGTGTTAGCCAGGAGCCTGTTCTTTTGGCACACGATTTTATAAGAATATTAGGCAC 1581

Qy 450 Gly-----ArgAspAspValThrAspGluGluMetGlu 460
Db 1582 GGTCTCATCGGCACAAAGTACGAGAATGAATCCGAGGATAAGGTCGGGAACATCATCGAG 1641
Qy 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
Db 1642 AACGGGGCAAAATGGCGAATGCTCATGACTTTATTACTGCTTGGCTGAAGGTTATGAG 1701
Qy 481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
Db 1702 ACCAATGTTGGCAGCGTGGCTTTCTCTTTCAGGTGGCCAGAAACAGCGCATTTGCAATC 1761
Qy 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
Db 1762 GCCCGTGGCTGTTAGTGACCAAAATCTGCTCTCGATGAAGCTACTTTCGGCGCTTG 1821
Qy 521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThr 540
Db 1822 GACAAATCTCCGAGGCGTGTTCAGCAGCTTTGGAGGGCGAGCTGAAGCCGAACT 1881
Qy 541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560
Db 1882 ACTATTGTGATCGCTCATCGCTTTCCAGCATCAAAACCGCGCACACATTTGGTTCG 1941
Qy 561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580
Db 1942 GTCAATGGCAAAATGCTGAACAAGCACTCAGCATGAATTTGTTGACCGCGAGCGCT 2001
Qy 581 TyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMet 600
Db 2002 TATCGCAAACTTGTGAGGCTCAACGATATCAATGAACAAGCAAGCAAGCTCACGCCCTGGAG 2061
Qy 601 ThrTyrSerThrGluArgLysThrAsnSer-----LeuProLeuHisSerValLysSer 618
Db 2062 GAGCGGACGCTGAGGATCTCAGATGCGATATGTCAGATATGTCAAAATCAAAATCGGTCAAGC 2121
Qy 619 IleLysSerAspPhe-----IleAspLys-----AlaGluGluSer 630
Db 2122 GCATCATCGATCTCGACGGAAAAACCCACACCAATGTCGCGACGGGACCCACCAAGTCT 2181
Qy 631 ThrGlnSerLysGluIleSer-----LeuProGluValSerLeuLeu 644
Db 2182 GTTTCAGCGCGATTTCTTTCTAAAGACCCCCCGAAACAACTCCGAAATCACTCATATGG 2241
Qy 645 LysIleLeuLys-----LeuAsnLysProGluTrpProPheValValLeuGly 660
Db 2242 AGCTGCTCAAAATTTGCTCTTCAACCGCTTGAATCCCGTACATGTCATCGGT 2301
Qy 661 ThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLys 680
Db 2302 CTTGTCTTCTCAGTGTAGTGTGTGGTGGCCAAACCCACCAAGCAGTGTATATGCTAAA 2361
Qy 681 IleIleThrMetPheGly-----AsnAsnAspLysThrThrLeuLysHisAspAlaGlu 698
Db 2362 GCCATCAGCACACTCTCGCTCCCAAGATCACAATATAGCAAGCTTCGACATGATCGGAT 2421
Qy 699 IleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGln 718
Db 2422 TTTGTCATATGATGTTCTTCGTTGGTGTATCATTCAGTTTATCAGCAGTCAACCAAT 2481
Qy 719 GlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPhe 738
Db 2482 GGTGCTGCATTTGCCGTATGCTCCGAGAGACTTATTCGTCGCGGAGAGACGCTGCTTT 2541
Qy 739 LysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGly 758
Db 2542 CGGACGATACTCGTCAAGACATTTGCTTTTTCAGAGAAAGAGATAGCACCGCGCT 2601
Qy 759 LeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIle 778
Db 2602 CTGACCTCTTCTCTCCACCGAGACGATCTCTCCGGTGTAGCGGTGTGACTCTA 2661


```
QY 779 GlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyr 798
Db 2662 GGCACGATCTTGATGACCTCCACGACCTAGGAGCGCTATCATTTGCGCTGGCGATT 2721
QY 799 GlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMet 818
Db 2722 GGGTGGAAATTTGGCTTAGTTGTATCTCGGTTGTGCGGTTCTCTCGGCATCGGTTTC 2781
QY 819 IleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAla 838
Db 2782 TACGATTCATATGCTAGCCCGATTTCAATCACGCTCCAAAGCTTGCTATGAGGATCT 2841
QY 839 GlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGlu 858
Db 2842 GCAAACTTTCCTTCGAGGGCTACATCGTCTATCCGACAGTTGCGTCATTAACCCGGGAA 2901
QY 859 LysAlaPheGluGlnMetTyrGluMetLeuGlnThrGlnHisArgAsnThrSerLys 878
Db 2902 AGGATGCTCGGAGATTACCATGCCAGCTTGACGCACAAGCAGCAGCAGCTTAATC 2961
QY 879 LysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyr 898
Db 2962 TCTGCTTGGAGTCATCCCTGTTATATGCTGCTGCGCAGCACTTGTTCTTCTCGCTT 3021
QY 899 AlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGly 918
Db 3022 CGCTCGGGTTTGGTAGGAGGACACTTCTTGGTCACCAGCATATGACATTTCCGC 3081
QY 919 MetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuVal 938
Db 3082 TTCTTTGTTGTTCTCCGAGATTCTCTTGGTGCTCAATCCGCGGCGCACCGCTTTTCC 3141
QY 939 LeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuGlu 958
Db 3142 TTGACACACATGCGCAAGCGCGAAGAAATGCGCGCCCGCAATTCGACACTGTTGAC 3201
QY 959 LysLysProAsnIleAspSerArgSerGlnGluLysLysProAspThrCysGluGly 978
Db 3202 CGAAGGCCACAAATTGATTAATGCTGTGAGAGGGCGGAGAGCTCGAAACGGTGGAGGT 3261
QY 979 AsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeu 998
Db 3262 GAAATCGAATTTAGGAACGTGCACCTTCAGATACCCGACCGCGCCGACAGCCCTGCTCG 3321
QY 999 ArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGly 1018
Db 3322 CGCGGCTTGGACCTGACCGTGAAGCCGGAACAATGTTGCGCTTGTGCGACCCGCGGT 3381
QY 1019 CysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGln 1038
Db 3382 TGTGGCAAGAGTACACCATTTGCATTGCTTGAGCGCTTTTACGATGCGATTGCGGGTCC 3441
QY 1039 ValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIle 1058
Db 3442 ATCCTTGTGTGAGGAGGACATAAGTAATAACTAAATATCAACTCTACCAGCCTTTCTG 3501
QY 1059 AlaIleValProGlnGluProValPheAsnCysSerIleAlaGluAsnIleAlaTyr 1078
Db 3502 TCACGTGTACGAGGAGCCGACACTGTACCAGGCGCCATCAAGGAAAAACATCTTACTT 3561
QY 1079 GlyAspAsnSerArgValProLeuAspGluIleLysGluAlaAlaAsnAlaAsn 1098
Db 3562 GGTATTGCGAAGATGAGTACCGAAGAAATCTTGATTAGGCTTGCAAGGACGCTAAT 3621
QY 1099 IleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGly 1118
Db 3622 ATCTACGACTTCATCATGTCCTCCGAGGGCTTTAATACAGTTGTTGGCAGCAGGGA 3681
QY 1119 AlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLys 1138
Db 3682 GGCATGTGTGTCGCGGCAAAAGCAACGTGTGCGCAATGCGCGACCTTCTTTCGGAT 3741
QY 1139 ProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspAsnSerSerGluLysVal 1158
```

```
Db 3742 CCCAAAATCTTCTTCTTCGATGAAGCGAGCTCAGCCCTCGACTCCGAGTCAGAAAAGGTC 3801
QY 1159 ValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArg 1178
Db 3802 GTCCAGCGCGCTTTGATGCCCTGCCGAGGCCGACCAATCGCCGTTGCACACCGA 3861
QY 1179 LeuSerAlaIleGlnAsnAlaAspLeuIleValLeuHisAsnGlyLysIleLysGlu 1198
Db 3862 CTCAGCACGATTCAAAAGGCGGACGTTATCTATGTTTCGACCAAGGCAAGATCGTCGAA 3921
QY 1199 GlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAla 1218
Db 3922 AGCGGAACGCACAGCGAACTGGTCCAGAAAAAGGCGCGTACTACGAGCTGGTCACTTG 3981
QY 1219 GlnSerVal 1221
Db 3982 CAGAGCTTG 3990
```

Search completed: March 31, 2003, 15:32:18
Job time : 455.956 secs

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 30, 2003, 03:22:44 ; Search time 8405.14 Seconds
(without alignments)
2354.618 Million cell updates/sec

Title: US-09-873-409-5
Perfect score: 6159
Sequence: 1 MITGLIASLVNGACILPLML.....QELLNRDIYFKLVNAQSVQ 1222

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DEV=xlh
-Q=/cpn2.1/USPTO.spool/US9873409/runat_27032003.115420.19240/app_query.fasta_1.7544
-DB=EST-QPMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US9873409 @CCN 1.1 30544 @runat 27032003.115420.19240 -NCPU=6 -ICPU=3
-NO XLPXY -NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1092	17.7	669	13	BM013981	BM013981 603639619
2	1051	17.1	1033	14	BM926413	BM926413 AGENCOURT
3	1008	16.4	943	9	AL520322	AL520322 AL520322
4	995	16.2	760	13	BM016204	BM016204 603642859
5	918	14.9	545	9	AL040762	AL040762 DKF2P434C
6	903	14.7	929	14	BQ882401	BQ882401 AGENCOURT
7	856.5	13.9	948	13	BM471690	BM471690 AGENCOURT
8	834	13.5	998	14	BM904842	BM904842 AGENCOURT
9	833.5	13.5	2676	11	AK014319	AK014319 Mus muscu
10	777.1	12.5	894	12	BF584668	BF584668 602098406
11	770.5	12.5	1019	12	BG248052	BG248052 602359987
12	753	12.2	726	12	BQ293345	BQ293345 602390738
13	752.5	12.2	871	17	AZ682350	AZ682350 ENTKB16TF
14	751	12.2	944	12	BF796582	BF796582 602258463
15	744.5	12.1	872	11	AK020318	AK020318 Mus muscu
16	743.5	12.1	886	17	BH133685	BH133685 ENTNA47TR
17	734.5	11.9	932	17	AZ670821	AZ670821 ENTJN69TF
18	728.5	11.8	947	17	AZ683753	AZ683753 ENTIL96TF
19	726.5	11.8	886	17	AZ540627	AZ540627 ENTQ18TF
20	723.5	11.7	880	17	AZ687805	AZ687805 ENTLU52TF
21	721.5	11.7	939	14	BQ720763	BQ720763 AGENCOURT
22	720.5	11.7	897	17	AZ541090	AZ541090 ENTDS67TR
23	719	11.7	780	12	BG587938	BG587938 EST489713
24	711.5	11.6	913	17	BH155700	BH155700 ENTRO54TR
25	709.5	11.5	834	17	AZ548312	AZ548312 ENTRO54TR
26	708.5	11.5	853	17	AZ679807	AZ679807 ENTHI67TR
27	705.5	11.5	823	17	AZ533602	AZ533602 ENTCR50TR
28	697.5	11.3	861	14	BQ717101	BQ717101 AGENCOURT
29	692.5	11.2	1341	11	AY108285	AY108285 Zea mays
30	690.5	11.2	899	17	BH154857	BH154857 ENTRO47TF
31	690	11.2	835	13	BJ349604	BJ349604 BJ349604
32	689.5	11.2	891	17	AZ682250	AZ682250 ENTKT68TR
33	688	11.2	1123	11	AY108485	AY108485 Zea mays
34	687	11.2	715	14	BQ869512	BQ869512 QGD6G09.Y
35	684	11.1	820	13	BJ356540	BJ356540 BJ356540
36	683	11.1	600	13	BM486593	BM486593 pgm2n.pk0
37	682.5	11.1	919	17	AZ690701	AZ690701 ENTIY68TR
38	680	11.0	926	14	BQ123477	BQ123477 EST609053
39	680	11.0	1813	11	U66688	U66688 Homo sapien
40	677.5	11.0	795	12	BF133560	BF133560 601900192
41	674.5	11.0	815	13	BI409057	BI409057 602961127
42	670	10.9	852	17	BH720383	BH720383 BOH2294TF
43	669	10.9	547	13	BJ030711	BJ030711 BJ030711
44	667	10.8	937	17	BH150760	BH150760 ENTQ042TR
45	666	10.8	792	13	BJ352404	BJ352404 BJ352404

ALIGNMENTS

RESULT 1
BM013981
LOCUS
DEFINITION BM013981 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:5415556 5',
669 bp mRNA linear EST 30-OCT-2001
mRNA sequence.
ACCESSION BM013981
VERSION BM013981.1 GI:16528335
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 669)
AUTHORS NIH-MGC http://mgc.mci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL
COMMENT

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTp
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL2060 row: f column: 05
High quality sequence stop: 664.
Location/Qualifiers
1. .669
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5415556"
/clone_lib="NIH MGC 87"
/tissue_type="mammary adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: breast; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.383 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH MGC Library."
BASE COUNT 203 a 122 c 162 g 182 t
ORIGIN

Alignment Scores:

Pred. No.: 7,71e-109 Length: 669
Score: 1092.00 Matches: 215
Percent Similarity: 99.54% Conservative: 2
Best Local Similarity: 98.62% Mismatches: 0
Query Match: 17.73% Indels: 1
DB: 13 Gaps: 0

US-09-873-409-5 (1-1222) x BM013981 (1-669)

QY 292 AlaValPheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPhe 311
Db 17 AGTGTTCCTTAATGAATCAATAGCAGTATTGCATTGGAGCAGGCTCCCTCAGTTT 76
QY 312 GluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLys 331
Db 77 GAAACCTTCGCAATAGCCCGAGGAGCTGCTTCATATTTTCCAGGTATTGATAAGAAA 136
QY 332 ProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrVal 351
Db 137 CCCAGTATAGATAACTTTTCCACAGCTGGGATATAAACCTGAATCCATAGAACTGTG 196
QY 352 GluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGly 371
Db 197 GAATTTAAATAGTTCTTTCAATATTCATCCATCAGACCATCTATCAAGATTCGAAGGT 256
QY 372 LeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGly 391
Db 257 CTGAATCTCAGAAATTAAGTCGGAGACAGCTCGCTTGGTCTCAATGCGAGTGG 316
QY 392 LysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMet 411
Db 317 AAGAGTACGGTAGTCCAGCTTCTGACAGGTTATATGATCGGATGATGCTTTATCATG 376
QY 412 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisIleGlyVal 431
Db 377 GTGATGAGAAATGATCATCAGAGCTTTTAAATGTGGCATTTATCCAGACCATATTGGAGT 436
QY 432 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 451
Db 437 GTTAGTCAAGAGCTGTTTGTTCGGACCCACCATCATGTAACAATACTAAGTATGACGA 496
QY 452 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 471
Db 497 GATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAATGCGTATGATTTT 556

QY 472 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyAlaGlnMetSer 491
Db 557 ATCATGAGGTTTCCTAATAATTAATCATTTAGTGAGGAAAAAGAGCTCAATGAGT 616
QY 492 GlyGlyGlnLysGlnArg-IleAlaIleAlaArgAlaLeuValArgAsnPro 508
Db 617 GGAGGGCAGAAAACAGAGGATCGCAATTGCTGCTCTAGTTTCGAACCCC 668

RESULT 2

BM926413 1033 bp mRNA linear EST 12-MAR-2002
LOCUS AGENCOURT_6600788 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5764845
DEFINITION 5', mRNA sequence.

ACCESSION BM926413
VERSION BM926413.1 GI:19376792
KEYWORDS EST.
SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 1033)

AUTHORS

NIH-MGC http://mgi.nci.nih.gov/.

TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Life Technologies, Inc.

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLNL218 row: o column: 22

High quality sequence start: 3

High quality sequence stop: 686.

FEATURES
source

1. .1033
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5764845"
/clone_lib="NIH MGC_114"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: EcoRV (destroyed); RNA source anonymous pool of 6
male brains, age range 23-27 yo. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.5 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 019. Note:
this is a NIH_MGC Library."

BASE COUNT 289 a 212 c 218 g 312 t

ORIGIN

Alignment Scores:
Pred. No.: 5.31e-104 Length: 1033
Score: 1051.00 Matches: 221
Percent Similarity: 94.4% Conservative: 2
Best Local Similarity: 93.64% Mismatches: 13
Query Match: 17.06% Indels: 3
DB: 14 Gaps: 0

US-09-873-409-5 (1-1222) x BM926413 (1-1033)

QY 293 ValPheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGlu 312
Db 180 GTTTTCTTTAGTGAATCCATAGCAGTTATTGCATTGGAGCAGCTCCCTCACTTGA 239
QY 313 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysPro 332
Db 240 ACCTTCGAATAGCCCGAGGAGCTGCTTTTCATATTTTCCAGGTCAATGATAAGAACCC 299


```

QY 333 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 352
|
|
Db 300 AGTATAGATAACTTTTCCACAGCTGGATATAACCTGAATCCATAGAGAACTGTGGAA 359
|
|
QY 353 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLysGlyLeu 372
|
|
Db 360 TTTAAAAATGTTCTTTCAATATATCCATCAAGACCATCTATCAAGATTCGAAAGTCTG 419
|
|
QY 373 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 392
|
|
Db 420 AATCTCAGATTAAGTCTGGAGAGACAGTCGCTTGGTCGGTCTCAATGGCATGGGAAG 479
|
|
QY 393 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 412
|
|
Db 480 AGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCGGATGATGGCTTTATCATGGTG 539
|
|
QY 413 AspGluAsnAspIleArgAlaLeuLeuValArgHisTyrArgAspHisIleGlyValVal 432
|
|
Db 540 GATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTTATCGAGACCATATTTGGAGTGGTT 599
|
|
QY 433 SerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAsp 452
|
|
Db 600 AGTCAAGAGCCTGTTTGTTCGGGACACCATCAGTAACAAATATCAAGTATGGACGAGAT 659
|
|
QY 453 AspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIle 472
|
|
Db 660 GATGTGACTGATCAAGAGATGGAGAGCAGCAAGGAAGCAATCGGTATGATTTATC 719
|
|
QY 473 MetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGly 492
|
|
Db 720 ATGGAGTTTCTCTAATAA-TTTAATACATTGGTAGGGGAANAAGAGAGCTCCAATCAGTGA 778
|
|
QY 493 GlyGlnGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIle 512
|
|
Db 779 AGCCAGAACCG-AGGATGCCAATGTCTCTGCTTGTAGTTCGAAACCCAGGAT-CTGAAT 836
|
|
QY 513 LeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaVal 528
|
|
Db 837 TTATAGTAGACTAGCTGCTGCTCGCTGGATTCCAGAAACCCAGTCAACTGTGTC 884
|
|
RESULT 3
AL520322
LOCUS
DEFINITION AL520322 LTI_NFL004_NBC2 Homo sapiens cDNA clone CS0DB06YC15 5
prime, mRNA sequence.
ACCESSION AL520322
VERSION AL520322.1 GI:12783815
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 943)
AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
Bp 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
source
1..943
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DB06YC15"
/clone_lib="LTI_NFL004_NBC2"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH108"
/notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and

```

cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 253 a 225 c 253 g 211 t

ORIGIN

Alignment Scores:
 Pred. No.: 2,36e-99 Length: 943
 Score: 1008.00 Matches: 194
 Percent Similarity: 80.40% Conservative: 48
 Best Local Similarity: 64.45% Mismatches: 58
 Query Match: 16.37% Indels: 1
 DB: 9 Gaps: 0

US-09-873-409-5 (1-1222) x AL520322 (1-943)

```

QY 921 IleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAla 940
|
|
Db 5 TTATATTATTTTTCAGCTGTGCTTGTGGTCCATGGCGTGGGCAAGTCAGTTTCATTGCT 64
|
|
QY 941 ProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLys 960
|
|
Db 65 CCTGACTATGCCAAAGCCAAATATATCAGCAGGCCACATCATCATGATCATTTGAAAAAAC 124
|
|
QY 961 ProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeu 980
|
|
Db 125 CCTTGAATTGACAGCTACAGCAGCGAAGGCCTAATCCCGAACACATTTGGAAGAAATGTC 184
|
|
QY 981 GluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGly 1000
|
|
Db 185 ACATTTGGTGAAGTTGTATTCAACTATCCACCCGACCGACATCCAGTCTTCAGGA 244
|
|
QY 1001 LeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGly 1020
|
|
Db 245 CTGAGCTGGAGGTGAAGAGGCGCAGAGCTGGCTCTGTGGGCGAGCAGTGGCTGTGGG 304
|
|
QY 1021 LysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlnValLeu 1040
|
|
Db 305 AAGAGCACAGTGTGTGAGTCTCTGGAGCGGTTCTACGCCCTTGGCAGGGAAGTGTG 364
|
|
QY 1041 PheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIle 1060
|
|
Db 365 CTGTGAGCGCAAGAAATAAAGCGACTGAATGTTCACTGGCTCCGAGCACACCTGGGCATC 424
|
|
QY 1061 ValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAsp 1080
|
|
Db 425 GTGTCCAGAGGCCCATCTCTTTGACTGCGCATTTGCTGAGAACATTTGCTATGGAGAC 484
|
|
QY 1081 AsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHis 1100
|
|
Db 485 AACAGCCGGTGTGTGTACAGGAAGAGATCGTGAGGCGCAGCAAGAGGCGCAACATACAT 544
|
|
QY 1101 SerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGln 1120
|
|
Db 545 GCCTTCATCGATCACTGCTTAATAATATAGCACTAAGTAGAGACAAAGAACTCAG 604
|
|
QY 1121 LeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLys 1140
|
|
Db 605 CTCTCTGGTGGCCAGAAACAAACGCAATTCATAGCTCGTCCCTTTGTAGACAGCCCTCAT 664
|
|
QY 1141 IleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGln 1160
|
|
Db 665 ATTTTGTCTTTGGATGAAGCCACGTCAGCTCTGATACACAAAAGTAAAAAGTGTTCCAA 724
|
|
QY 1161 HisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSer 1180
|
|
Db 725 GAAGCCCTGGACAAAGCCAGAGAAAGCCGACCTGATGTGATGTCTACCCGCTGTC 784
|
|
QY 1181 AlaIleGlnAsnAlaAspLeuIleVal-ValLeuHisAsnGlyLysIleLysGluGlnGln 1200
|
|

```



```

Db 785 ACCATCCAGAAATGAGAGTCTAATAGTGTGTTTCAGAAATGGCAGAGTCAAGGAGCATGG 844
Qy 1200 yThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAlaGlnSe 1220
Db 845 CAGCATCAGAGCTGTCTGGCAGAAAGGCACTATTTTCAATGTCAGTCCAGGC 904
Qy 1220 x 1220
Db 905 T 905

RESULT 4
BM016204 760 bp mRNA linear EST 30-OCT-2001
DEFINITION 603642659F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:5418615 5',
mRNA sequence.
ACCESSION BM016204
VERSION BM016204.1 GI:16530558
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 760)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L1AM12068 row: e column: 16
High quality sequence stop: 740.
FEATURES
source
1..760
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5418615"
/clone_lib="NIH_MGC_87"
/tissue_type="mammary adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: breast; Vector: PCMV-SPORE6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.383 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT 229 a 150 c 187 g 193 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 4.27e-98 Length: 760
Score: 995.00 Matches: 203
Percent Similarity: 98.07% Conservative: 0
Best Local Similarity: 98.07% Mismatches: 2
Query Match: 16.16% Indels: 2
DB: 13 Gaps: 0

US-09-873-409-5 (1-1222) x BM016204 (1-760)
Qy 330 LysLysProSerIleAspAsnPhSerThrAlaGlyTyrLysProGluSerIleGluGly 349
Db 3 AAGAAACCCAGTATAGATAACTTTTCCACAGCTGGATATAAACCCTGAATCCATAGAGGA 62
Qy 350 ThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeu 369
Db 63 ACTGTGAATTTAAAGATGTTCTTCAATTATCATCATAGACCATCTATCAAGATTCG 122
Qy 370 LysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGly 389

```

```

Db 123 AAAGGTCGAATCTCAAAATTAAGTCTGGAGAGACAGTCCGCTTGGTTCGATCGC 182
Qy 390 SerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspGlyPhe 409
Db 183 AGTGGGAAGAGTAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCGGATGATGCGCTTT 242
Qy 410 IleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIle 429
Db 243 ATCATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATN 302
Qy 430 GlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyr 449
Db 303 GGAGTGTAGTCAAGAGCGCTGTTTGTTCGGACACCATCATAGTAACAATATCAAGTAT 362
Qy 450 GlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyr 469
Db 363 GGACGAGATGATGTGACTGATGAAGAGATGGAGAGCAGCAGGAGCAAGCAATCGGTAT 422
Qy 470 AspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGln 489
Db 423 GATTTTATCATGAGTTCCTAATAAATTTAATACATTTGGTAGGGGAAAAAGAGCTCAA 482
Qy 490 MetSerGlyGlnLysGlnArgIleAlaIle-AlaArgAlaLeuValArgAsnProLys 509
Db 483 ATGAGTGGAGGGCAGAAACAGAGGATCGCAATTGGCTCGTCCCTTAGTTTCGAAACCCCAA 542
Qy 509 sIleLeuIleLeu-AspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValG 529
Db 543 GATTCGTGATTTTCAGATGAGGCTACGTCGCCCTGGATTCAAGAAAGCAAGTCAGCTGTC 602
Qy 529 InAlaAlaLeuGluLys 534
Db 603 AAGCTGCACCTGGAGAG 619

RESULT 5
AL040762 545 bp mRNA linear EST 29-FEB-2000
LOCUS DKFZp434C1815 r1 434 (synonym: htes3) Homo sapiens cDNA clone
DEFINITION DKFZp434C1815 5', mRNA sequence.
ACCESSION AL040762.1 GI:5409708
VERSION AL040762
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 545)
AUTHORS Bloecker,H., Boecker,M., Brandt,P., Mewes,H.W., Gassenhuber,J. and
Wiemann,S.
TITLE EST (Bloecker, et al.)
JOURNAL Unpublished (1999)
COMMENT Contact: Bloecker H
MIPS
Am Klopferspitz 18a D-82152 Martinsried, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by GBF (National Research Centre for Biotechnology Ltd.,
Braunschweig/Germany) within the cDNA sequencing consortium of the
German Genome Project.
s1 sequence also available.
This clone (DKFZp434C1815) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
source
1..545
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZp434C1815"
/clone_lib="434 (synonym: htes3)"
/tissue_type="testis"
/dev_stage="adult"

```



```
/lab_host="DH10B"
/Note="Vector: pSport1; Site 1: NotI; Site 2: SalI"
BASE COUNT 171 a 89 c 138 g 147 t
ORIGIN

Alignment Scores:
Pred. No.: 6,63e-90 Length: 545
Score: 918.00 Matches: 181
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 14.91% Indels: 0
DB: 9 Gaps: 0

US-09-873-409-5 (1-1222) x AL040762 (1-545)

QY 333 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 352
DB 3 AGTATAGATAACTTTTCCACAGCTGGATATAACTGATCCATAGAGAACTGTGGAA 62
QY 353 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu 372
DB 63 TTTAAAAATGTTTCTTCAATATCCATCAAGACCATCTATCAAGATTCTGAAAGGCTCTG 122
QY 373 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 392
DB 123 AATCTCAGAAATTAAGTCTCGAGAGACAGTCGCTTGGTCTCAATGCGAGTGGGAAG 182
QY 393 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 412
DB 183 AGTACCGGTAGTCAGCTTCTCGACAGGTTATATGATCCGGATGATGGCTTTATCATGGTG 242
QY 413 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal 432
DB 243 GATGAGAATGACATCAGAGCTTTTAAATGTGCGGCATTTATCGAGACCATATTGAGTGGTT 302
QY 433 SerGlnGluProValLeuPheGlyThrIleSerAsnAsnIleLysTyrGlyArgAsp 452
DB 303 AGTCAAGAGCGCTGTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGCGACGAGAT 362
QY 453 AspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIle 472
DB 363 GATGTGACTGATGAAGATGGAGAGACGACGACGAGGAAAGCAATCGGTATGATTATTAIC 422
QY 473 MetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGlyAlaGlnMetSerGly 492
DB 423 ATGAGTTTCTTAATAATTTATACATTGTTAGGTAGGGGAAAAAGAGCTCAATGAGTGA 482
QY 493 GlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIle 512
DB 483 GGGCAGAAACAGAGGATCGCAATTGCTCGTCTTGTAGTTTCGAAACCCCAAGATTCTGATT 542
QY 513 Leu 513
DB 543 TTA 545

RESULT 6
BO882401
LOCUS
DEFINITION AGENCOURT_8627902 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:6291782
5', mRNA sequence.
ACCESSION BO882401
VERSION BO882401.1 GI:22274409
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 929)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
```

```
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2492 row: c column: 15
High quality sequence stop: 677.
Location/Qualifiers
1. 929
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6291782"
/clone_lib="NIH_MGC_43"
/tissue_type="normal pigmented retinal epithelium"
/lab_host="DH10B (phage-resistant)"
/Note="Organ: eye; Vector: pOT87; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
BASE COUNT 281 a 176 c 237 g 234 t
ORIGIN

Alignment Scores:
Pred. No.: 7,64e-88 Length: 929
Score: 903.00 Matches: 188
Percent Similarity: 73.02% Conservative: 61
Best Local Similarity: 55.13% Mismatches: 58
Query Match: 14.66% Indels: 34
DB: 14 Gaps: 3

US-09-873-409-5 (1-1222) x BO882401 (1-929)

QY 176 AlaTyrSerLysAlaGlyAlaValAlaGluValLeuSerSerIleArgThrValI1 195
DB 1 GCGTATGCAAAAGCTGGAGCAGTACTGGAAGAGTCTTGGCAGCAATTAGACTGGTAT 60
QY 195 eAlaPheArgAlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTy 215
DB 61 TGCATTGGAGGACAAAGAAAGAACTTGAA----- 91
QY 215 xAlaTrpPheTyrPheProGlnTrpLeuLeuSerCysValLeu**PheValArgTyrTh 235
DB 92 -----AGGTACAA 99
QY 235 rGlnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSe 255
DB 100 CAAAATTTAGAAAGCTTAAAGAAATTGGGATAAAGAGCTATTACAGCCAATATTTC 159
QY 255 rLeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyTh 275
DB 160 TATAGTGTGCTTCTCTCTGCTATCATCTATGCTATGCTCTGCTCTGCTCTGCTATGGGAC 219
QY 275 rSerLeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePh 295
DB 220 CACCTTGGTCTCTCAGGGAA-----TATTCTATGGACAGTACTCACTGATTCCT 273
QY 295 eSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAl 315
DB 274 TTTCTGTTAATTGGGGCTTTTAGTGTGGACAGGACTCTCCAAGCATTTGAAGCAATTTC 333
QY 315 aIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAs 335
DB 334 AAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATAATTGATAAAGCAAGTATTGA 393
QY 335 pAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAs 355
DB 394 CAGCTATTGCAAGAGTGGGCAACCAACAGATAATAATTAAGGGAAATTTGGAAATTCAGAA 453
```

Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2492 row: c column: 15
High quality sequence stop: 677.
Location/Qualifiers
1. 929
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6291782"
/clone_lib="NIH_MGC_43"
/tissue_type="normal pigmented retinal epithelium"
/lab_host="DH10B (phage-resistant)"
/Note="Organ: eye; Vector: pOT87; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."


```

QY 355 nValSerPheAsnTyrProSerArgProSerIleLysLeuLysGlyLeuAsnLeuAr 375
Db 454 TGTTCACTTACCTTACCACTCGAAGAAGATTAGATCTTGAAGGCTCTGAACCTGAA 513
QY 375 gileLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrVa 395
Db 514 GGTGCAGATGGGCAGACGGTGGCCCTGGTTGGAACACAGTGGCTGTGGGAAGACACAAC 573
QY 395 lValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAs 415
Db 574 AGTCCAGCTGATGCAGAGGCTCTATGACCCACAGAGGGGATGGTCAGTGTTCATGGACA 633
QY 415 nAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGl 435
Db 634 GGATATTAGACCAATAATTAAGTTCTTACGGGAATCATTTGGTGTGGTGCAGGA 693
QY 435 uProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspValTh 455
Db 694 ACCTGTATTGTTGCCACCAAGATAGTGAACACATTCCTATGGCCGTGAAATGTCCAC 753
QY 455 rAspGluGluMetGluArgAlaAlaArgGluAlaAsnAla-TyrAspPheIleMetClup 475
Db 754 CATGGATGCAATTGAGAAGCTGTCAGGAAGCAATGCCCTATGACTTTATCATGAAC 813
QY 475 heProAsnLysPheAsnThrLeuValGlyGlyLysGlyAlaGlnMetSerGlyGly---G 494
Db 814 TGCTCTATAAATTGACACCTGTTGGANAAGAGGGGCCAATTGAGTGTGGGGCAC 873
QY 494 lnLysGlnArgIleAlaAlaAlaArg-AlaLeuVal-ArgAsnProLysIleLeu 511
Db 874 AAACCAAGGATCTCCCTGGCAGCTGGCCCTGGTTCCCAACCCCAAGATCCTC 928

RESULT 7
BM471690 BM471690 948 bp mRNA linear EST 05-FEB-2002
LOCUS AGENCOURT_6465349 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5539117
DEFINITION 5', mRNA sequence.
ACCESSION BM471690
VERSION BM471690.1 GI:18520732
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 948)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: csapbs-r@mail.nih.gov
Tissue Procurement: ATCC/DCTP/DTP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM12233 row: b column: 14
High quality sequence stop: 569.
Location/Qualifiers
1. 948
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5539117"
/clone_lib="NIH_MGC_72"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2 kb. Library constructed by Life
Technologies."
BASE COUNT 285 a 187 c 194 g 282 t

```

ORIGIN

Alignment Scores:
 Pred. No.: 1e-82 Length: 948
 Score: 856.50 Matches: 186
 Percent Similarity: 89.67% Conservative: 5
 Best Local Similarity: 87.32% Mismatches: 14
 Query Match: 13.91% Indels: 8
 DB: 13 Gaps: 1

US-09-873-409-5 (1-1222) x BM471690 (1-948)

```

QY 293 ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGlu 312
Db 205 GTTTCTTTTAGTGAATCCATAGCAGTTATTGTCATGGAGCAGCAGTCCCTCACTTTGAA 264
QY 313 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPro 332
Db 265 ACCTTCGCAATAGCCGAGGAGCTGCTTTTCATATTTCAGGTTATTGATAAGAACCC 324
QY 333 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 352
Db 325 AGTATAGATAACTTTTCCACAGCTGGATATAACCTGAATCCATAGAGAACTGTGGAA 384
QY 353 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu 372
Db 385 TTTTAAAAATGTTTCTTTCAATTATCCATCAAGACCATCTATCAAGATTCTGAAAGGCTCG 444
QY 373 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 392
Db 445 AATCTCAGAAATTAAGTCTGGAGACAGTCGCTTGGTTCGATCAATGGCAGTGGGAAG 504
QY 393 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 412
Db 505 AGTACGGTAGTCAGCTTCTGCAGAGTTATATGATCCGGATGATGGCTTATCATGGTG 564
QY 413 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIle-GlyValVa 432
Db 565 GATGAGAATGACATCAGAGCTTTTAAATGTGGGCAATTATCGAGACCATATTTGGAGTGGT 624
QY 432 lSerGlnGluPro--ValLeuPheGlyThrThrIleSerAsnAsnIleLysTyr-GlyAr 451
Db 625 TAGTCAAGAGCCCTGGTTTGTTCGGGACCACCATCAGTAACAAATATCAAGTATGGGACG 684
QY 451 gAspAspValThrAspGluGlu-MetGluArgAlaAlaArgGluAla-AsnAlaTyrAsp 470
Db 685 AGATGATGGAGCTGATGAAACAAATGGAACACACAGGGAAACCAAGGCGAATGAA 744
QY 471 PheIleMetGlu-PheProAsnLysPheAsnThrLeuVal---GlyGluLysGlyAlaGl 489
Db 745 TTTATCATGGAATTTCCCAAAAAAATTTAATACATTGGAAAGGGGAAAAAGGACCTTCA 804
QY 489 nMetSerGlyGlyGlnLysGlnArg 497
Db 805 AGGAATGGAAGGCCCAAAACCCCAAG 829

```

RESULT 8

BM904842 BM904842 998 bp mRNA linear EST 12-MAR-2002
 LOCUS AGENCOURT_6699581 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5557655
 DEFINITION 5', mRNA sequence.
 ACCESSION BM904842
 VERSION BM904842.1 GI:19355221
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 998)
 NIH-MGC http://mgs.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.


```

http://image.llnl.gov
Plate: LLAM12279 row: f column: 24
High quality sequence stop: 738.
Location/Qualifiers
1. .998
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5557655"
/clone_lib="NIH MGC 72"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; vector: pCMV-SPORT6; Site_1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr.
Average insert size 2 kb. Library constructed by life
Technologies"

```

Alignment Scores:		
Pred. No.:	3,24e-80	Length:
Score:	834.00	Matches:
Percent Similarity:	74.12%	Conservative:
Best Local Similarity:	52.40%	Mismatches:
Query Match:	13.54%	Indels:
DB:	14	Gaps:
		0
US-09-873-409-5 (1-1222)	x	BM904842 (1-998)

Qy	745	AspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAla	764
Db	61	CATGTGATGTGGTTTGATGCCCTAAAAACACCACTGGAGCAATTCATACAGGCTCGCC	120
Qy	765	IleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsn	784
Db	121	AATGATGCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTAATTACCCAGAAT	180
Qy	785	AlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyArgIleTyGluMetThrPhe	804
Db	181	ATAGCAAACTCTGGACAGGAAGTAATATATCTTCATCTATGGTTGGCAACTTAACACTG	240
Qy	805	LeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMet	824
Db	241	TTACTCTTAGCAATTGTACCCTCATCTTGCATATAGCAGGAGTGTGTGAAATGAAATGTTG	300
Qy	825	ThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGlu	844
Db	301	TCTGGCAAGCACTGAAAGATAAAGAAAGAACTAGCAAGTGCTGGGAAGATCGCTACTGAA	360
Qy	845	AlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMet	864
Db	361	GCAATAGAAAACTTCCGAACCGTTGTTCTTTTGACTCAGGAGCAGAAAGTTTGAACATATG	420
Qy	865	TyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGly	884
Db	421	TATGCTCAGAGTTTGCAGGTACCATACAGAAACTCTTTGAGGAAGCACACATCTTTTGA	480
Qy	885	SerCysTyrAlaPheSerHisAlaPheIleTyArgPheAlaTyAlaAlaGlyPheArgPhe	904
Db	481	ATTACATTTTCTTCACCAGGCAATGATGATTTTTTCTATGCTGGATGTTTCCGGTTT	540
Qy	905	GlyAlaTyLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThr	924
Db	541	GGAGCCTACTTGGTGGCACATAAACTCATGAGCTTTGAGGATGTTCTGTTAGTATTTTCA	600
Qy	925	AlaIleAlaTyGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTySer	944

Db	601	GCTGTTGCTTTTGGTGCCATGGCCGTCGGGCAAGTCAGTTCATTCTTGTCTCGACTATGCC	660
Qy	945	LysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnLeaSp	964
Db	661	AAAGCCAAAATATCAGCAGCCACCATCATCATGATCATTTGAAAAAACCCTTTTGTATTGAC	720
Qy	965	SerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuLeuPheArgGlu	984
Db	721	AGCTACAGCAGCGAAGGCCTTAATGCCGAACACATTTGGAAGGAATGTCACATTTGGGGTGA	780
Qy	985	ValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSe	1004
Db	781	GTGTATTCAACTATCCACCCGACCGACATCCAGTCTTCAGGAGCTCAGCCCTGGA	840
Qy	1004	rIleGluArgGlyLysThrValAlaPheVal-GlySerSerGlyCysGlyLysSerThrS	1024
Db	841	GGTGAAGAAGGGGAGACCCCTGGCTCTGTGGGGCAGCAGTGGCTGTGGGAAGACACAG	900
Qy	1024	er-ValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGly	1043
Db	901	TGGGTGCAGTCTCTGGAGCGGTCTCAACCCCTTGGCAAGGAATGCTCTGTGATGCC	960
Qy	1044	ValAspAlaLys-GluLeuAsnValGlnTnp	1053
Db	961	AAAGAATAAACCGATGGATGTTCATGG	991
RESULT	9	AK014319	2676 bp mRNA linear HTC 19-JAN-2002
LOCUS	AK014319		
DEFINITION		Mus musculus 14, 17 days embryo head cDNA, RIKEN full-length enriched library, clone:322401P09:ATP-binding cassette, sub-family B (MDR/TRAP), member 8, full insert sequence.	
ACCESSION		AK014319	
VERSION		AK014319.1	GI:112852089
KEYWORDS		HTC; CAP trapper.	
SOURCE		Mus musculus (strain: C57BL/6J) 14, 17 days embryo head cDNA to mRNA, clone lib: RIKEN full-length enriched mouse cDNA library	
		Clone: 3222401P09.	
ORGANISM		Mus musculus	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;	
AUTHORS		1 Carninci, P. and Hayashizaki, Y.	
TITLE		High-efficiency full-length cDNA cloning	
JOURNAL		Meth. Enzymol. 303, 19-44 (1999)	
MEDLINE		99279253	
PUBMED		10349636	
REFERENCE		2	
AUTHORS		Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.	
TITLE		Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes	
JOURNAL		Genome Res. 10 (10), 1617-1630 (2000)	
MEDLINE		20499374	
PUBMED		11042159	
REFERENCE		3	
AUTHORS		Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kiteunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yanamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanishi, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.	
TITLE		RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer	
JOURNAL		Genome Res. 10 (11), 1757-1771 (2000)	
MEDLINE		20530913	
PUBMED		11076961	
REFERENCE		4	
AUTHORS		Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Konno, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, K.	

Db 897 CTTACCTGATGCTGGCTGTCACACCGCCCTCATGGAGTGGCCACCTGATGGGC 956
 Qy 164 ArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaVal 183
 Db 957 TCAGGCTCCGAAGCTGCTCCAGTGTCCAGGACAGATTGCCAGGCGACAGGTGA 1016
 Qy 184 AlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGlu 203
 Db 1017 GCAGATGAGGCCCTTGCCCAATGTTCCGACTGTGCGGCGCTTCGCGATGAGAGAGGAG 1076
 Qy 204 LeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTyrPheTyrPheProGlnTyr 223
 Db 1077 GAGGAACGC-----TATCAAGCAGAA 1097
 Qy 224 LeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaTyr 243
 Db 1098 CTGAGTCTATGCTGCT-----AAGCAGAGAA 1127
 Qy 244 PheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhe----- 261
 Db 1128 CTGGGC-----AGGGCATCGCTTGTTCAGAGGCTCTCCAACATCGCTTTCAACTGT 1181
 Qy 262 PheMetAsnGlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuLeuAsnGly 281
 Db 1182 ATGGTCTTGGGCACC-----CTGTTTATGGGGCTCCCTTGTGGCTGGACAG 1229
 Qy 282 Glu---ProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
 Db 1230 CAGCTGAAGGGGACCTCATGCTCTCTGGTGGCTTCCAGACAGTACAGAGTCT 1289
 Qy 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGly--- 319
 Db 1290 ATGGCCAGCTCTCTCTCTGTTTGTGTAG-----GTGGTACGTGGGCTG 1334
 Qy 320 ---AlaAlaPheHisIlePheGlnValIleAspLysPheProSerIleAspAsnPheSer 338
 Db 1335 AGTGGCGGAGCCGAGTCTTCGAATACATGGCCCTGAGCCCTGTCATCCCA-----TTG 1388
 Qy 339 ThrAlaGlyTyr-----LysProGluSerIleGluGlyThrValGluPheLysAsnVal 356
 Db 1389 ACCGGGGTACTGTCATCCCAAGGACATTCGGTGTTCATACCTTCCAAATGTC 1448
 Qy 357 SerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIle 376
 Db 1449 ACCTTCAGTACCCCTGCAGACCTGCTCAATGCTCAAGGACTTCACCTGGAAGCTG 1508
 Qy 377 LysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValVal 396
 Db 1509 CCCTCTGGCAAGATTGGCTCTTGTGGCCAGTCTGGGGGAGGAAGACCAAGTTGCC 1568
 Qy 397 GlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAsp 416
 Db 1569 TCCTCTGGAACGCTTCTATGACCTTGAAGCTGGCTGGCTGAGCTGTGATGGGCAAGC 1628
 Qy 417 IleArgAlaLeuAsnValArgHisTyrArgAspHis---IleGlyValValSerGlnGlu 435
 Db 1629 CTGCGAATCTCAACCCCTCTGCTCCGGGCGGCGGCTCATAGTTTTCATCAGCCAGGAG 1688
 Qy 436 ProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspValThr 455
 Db 1689 CCAGTCTGTTTGCACCAACCATCATGGAGATATCCGATTGGGAAGCTGGATGCTTC 1748
 Qy 456 AspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPhe 475
 Db 1749 GATGAAGAGGTGPACAGCTGCACGAGAGCCAAATGCCCGAGTTCATCAGCAGCTTC 1808
 Qy 476 ProAsnLysPheAsnThrLeuValGlyGlyAlaGlnMetSerGlyGlyGlnLys 495
 Db 1809 CCGATGGCTACAGCACTGTGTTGTGTGAGCGGGGACACCTTGTCTGTGGCCAGAG 1868
 Qy 496 GlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuLeuAspGlu 515

Db 1869 CAGCGCTAGCCATCGCAGTGCCTCATCAAGCAGCCACAGTGTCTCTGACGAG 1928
 Qy 516 AlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAla 535
 Db 1929 GCCACCAAGTGTAGTGTGACAGATCCGAGAGGGTGTACAGAGGCGCTGGACCGGCC 1988
 Qy 536 SerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAsp 555
 Db 1989 AGTGTGCGCGCAGCGTGTGTGCTCATTCGCCACCGCTTAGTACTGTCTCGTGCAGCCAC 2048
 Qy 556 LeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMet 575
 Db 2049 TCATCATCTGTATGCCCAATGCCAAGTCTGTGAGCTGGAGCCACCAAGAACTCTT 2108
 Qy 576 AlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIle 590
 Db 2109 AAAAAAGGGGGCTCTATTTCAGAGCTTATCCGAGACAAACCTGT 2153
 RESULT 10
 BF584668 894 bp mRNA linear EST 12-DEC-2000
 LOCUS 602098406F1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4218385 5',
 DEFINITION mRNA sequence.
 ACCESSION BF584668
 VERSION BF584668.1 GI:11658386
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 894)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 Cloning Distribution: Incyte Genomics, Inc.
 Cloned through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM9798 row: 1 column: 02
 High quality sequence stop: 651.
 FEATURES
 source 1..894
 /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:4218385"
 /lab_host="NCI CGAP Co24"
 /note="Organ: colon; Vector: pCMV-SPORT6; Site: 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr.
 Average insert size 1.6 kb. Constructed by Life
 Technologies. Note: this is a NCI CGAP Library."
 BASE COUNT 263 a 203 c 242 g 185 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 2,18e-73 Length: 894
 Score: 771.00 Matches: 171
 Percent Similarity: 79.72% Conservative: 57
 Best Local Similarity: 59.79% Mismatches: 48
 Query Match: 12.52% Indels: 15
 DB: 12 Gaps: 3
 US-09-873-409-5 (1-1222) x BF584668 (1-894)
 Qy 293 ValPhePheSerValIleHisSerTyrCysIleGlyAlaAlaValProHisPheGlu 312
 Db 5 GTCTTCTTTTCCGTGTTAT-GGAGCAATTTCAGTGTGGACAGGCATCTCCAAATATTGAA 63

QY 313 ThrPheAlaIleAlaAaGlyAlaAlaPheHisIlePheGlnValIleAspTyrLysPro 332
 Db 64 GCCTTCGCCAATGCACGAGGAGCGCTTATGAAGTCTTCAAAATAATGATATAAGCCC 123
 QY 333 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 352
 Db 124 AGTATACACAGCTTCTCAAGAGTGGGCAAAACACGACACATACAGAAATCTGGAA 183
 QY 353 PheIleAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu 372
 Db 184 TTTAAGAATATTCACCTTCAGTTACCCCATCTCGAAAGAGTTCAGATCTTGAAGGGCCTC 243
 QY 373 AsnLeuArgIleLysSerGlyThrValAlaLeuValGlyLeuAsnGlySerGlyLys 392
 Db 244 AATCTGAAGTGAAGAGCGACAGCGTGGCTGGTGGCAACAGTGGCTGGGAAA 303
 QY 393 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 412
 Db 304 AGCACAACTGTCCAGCTGATGCAAGGCTCTAGACCCCTAGATGGCATGGTCAGTATC 363
 QY 413 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal 432
 Db 364 GACGGACAGACATCAGAACCATATGTGAGGTATCTGAGGGAGATCATGTGTGGTG 423
 QY 433 SerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAsp 452
 Db 424 AGTCAGAACTGTGCTGTTTGGCCACCATCGCCGAGAACATTCGCTATGGCCGAGAA 483
 QY 453 AspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIle 472
 Db 484 GATGTCAACCATGATGATGATGAGAAAGCTGTCAAGGAGCCCAATGCTATGACTTCATC 543
 QY 473 MetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGly 492
 Db 544 ATGAACCTGCCCACTTATGACCTGTGTGGTGGAGAGAGGGCGAGCTGAGTGGG 603
 QY 493 GlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIle 512
 Db 604 GGACAGAAACAGAGAAATCGC-ATTGCCCGGGCCCTGTGTCGCAATCCCAAGTGCCTTC- 661
 QY 513 LeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla-Le 532
 Db 662 GTGACCAAGCCACCTCAGC-CTGGATACAGAAAGTGAAGCTGGGGTTCAGGCGCCACCT 720
 QY 532 uGluLysAlaSerLysGlyArgThrIle-----ValValAlaHisArgIle 548
 Db 721 AGATAGGGTACAAAAGGCGGAGCCACCATTTGGGAGAGCTCATCGCTGTTCGTCGAAGCT 780
 QY 548 uSerThrIleArgSerAlaAspLeuIleVal-----ThrLeuLysAspGlyMetLeuAl 566
 Db 781 GACGGCCATCGGGCCGAGGGGGG-CTTGTGGGAGACACTCTGNAAGC-----TC 830
 QY 566 aGluLysGlyAlaHis 571
 Db 831 TGAAAAAGGTACTCAT 846
 RESULT 11
 BG248052
 LOCUS 602359987F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:4488404 5',
 DEFINITION mRNA sequence.
 ACCESSION BG248052.1 GI:12757867
 VERSION BG248052
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 1019)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgbbs-remail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM10334 row: n column: 21
 High quality sequence stop: 650.
 FEATURES
 source
 1..1019
 Location/Qualifiers
 /organism="Mus musculus"
 /strains="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:4488404"
 /clone_lib="NCI_CGAP_Mam1"
 /tissue_type="tumor, biopsy sample"
 /dev_stage="10 months, virgin"
 /lab_host="DH10B"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
 Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
 Library constructed by Life Technologies. Investigator
 providing samples: Gilbert Smith, NIH"
 BASE COUNT 276 a 262 c 283 g 196 t 2 others
 ORIGIN
 Alignment Scores:
 Pred. No.: 3,13e-73 Length: 1019
 Score: 770.50 Matches: 154
 Percent Similarity: 79.17% Conservative: 36
 Best Local Similarity: 64.17% Mismatches: 47
 Query Match: 12.51% Indels: 3
 DB: 12 Gaps: 1
 US-09-873-409-5 (1-1222) x BG248052 (1-1019)
 QY 977 GluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPhe 996
 Db 4 GAAGGAATGTGCAATTTAGTGGAGTCTGTTCAACTATCCACCCGACCCGATCCCA 63
 QY 997 IleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySer 1016
 Db 64 GTCTTTCAGGGGTGAGCTTGTAGGTGAAGAGGGCCAGACGCTGGCCCTGTGGGCGC 123
 QY 1017 SerGlyCysGlyLysSerThrSerValGlnLeuGlnArgLeuTyrAspProValGln 1036
 Db 124 AGTGGCTGCGGGAGAGCACAGTGGTCCAGCTGCTCGAGCGCTTCTACGACCCCATGGCT 183
 QY 1037 GlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSer 1056
 Db 184 GGATCAGTGTTCATGATGGCAAGAAATAAGCACTGNAATGTCCAGTGGTCCGAGCA 243
 QY 1057 GlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIle 1076
 Db 244 CAGCTGGCATTTGTGTCCCAAGAGCCATCTCTTTGACTGCAGCATCGCAGAGAACATT 303
 QY 1077 AlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAsnAla 1096
 Db 304 GCCTACGGAGACAAACAGCGGGTCTGTCTTATGAGGAGATTTGTAGGGGCGAGGAG 363
 QY 1097 AlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeu 1116
 Db 364 GCCAACATCCACCATTCATCGACTCGCTACTGTATATACACACAGAGTAGGAGAC 423
 QY 1117 LysGlyAlaGlnLeuSerGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeu 1136
 Db 424 AAAGGCACCTCAGCTGTGGGTGGGAGAGAGCGCATCGCCATCGCAGCGCCCTCGTC 483
 QY 1137 GlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnSerGlu 1156
 Db 484 AGACAGCCTCACATTTTACTTCTGGAGCAAGCAACATCAGCTCTGGATACAGAAAGTGA 543

Qy	1051	ValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCys	1070
Db	122	GTCCAGTGGCTCCGAGCTCAACTGGGCAATTGTGTCCAGGAACCACTTCTCTTTGACTGC	181
Qy	1071	SerIleAlaGluAsnIleAlaTrpGlyAspAsnSerAqValValProLeuAspGluIle	1090
Db	182	AGATCGCAGAGAACATCGCCTATGGAGACAACACCGGGTCTGTCCTCATGATGAGATT	241
Qy	1091	LysGluAlaAlaAsnAlaIleHisSerPheIleGluGlyLeuProGluLysTyr	1110
Db	242	GTGAGGCGACCAAGGAGGCCAACATCCACCCTTTCATCGAGACGCTGCCCAAAATAT	301
Qy	1111	AsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAla	1130
Db	302	AACACAAGAGTAGGAGACAAGGGGACGACAGCTCTCTGGGGGCCAGAACGACAGGATTGCC	361
Qy	1131	IleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAla	1150
Db	362	ATGCGCCGAGCCCTCATCAGACAGCTCGGTCTCTACTGCTGGATGAAGCCACGTCAGCT	421
Qy	1151	LeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArg	1170
Db	422	CTGGATACTCAGAGTCAAAAGGTGTTCAGGAAGCACTGGACAAGCCAGGAGGCGCGC	481
Qy	1171	ThrCysLeuValValThrHisArgIleSerAlaIleGlnAsnAlaAspLeuIleValVal	1190
Db	482	ACCTGATCCTGAGTCGCTCACCGCTGTCCACATCCAGAACCGGCACTTGATCGTGGT	541
Qy	1191	LleuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgA	1210
Db	542	GATTGACAAGCGCAAGGTCAAGGACACGCGCACCCACCACGAGCTGCTGGCGCAGAGG	601
Qy	1210	spIleTyrPheLysLeuValAsn	1217
Db	602	GCATCTATTCTCAACTGGTCAAC	625
RESULT	13		
LOCUS	AZ682350	871 bp DNA linear	GSS 14-DEC-
DEFINITION	ENTXB167F Entamoeba histolytica Sheared DNA Entamoeba histolytica genomic, DNA sequence.		
ACCESSION	AZ682350		
VERSION	AZ682350.1	GI:11819496	
KEYWORDS	GSS.		
SOURCE	Entamoeba histolytica.		
ORGANISM	Entamoeba histolytica		
REFERENCE	1 (bases 1 to 871)		
AUTHORS	Loftus B., Van Aken S. and Fraser, C.		
TITLE	Determination of clone end sequences from Entamoeba histolytica		
JOURNAL	HM1:IMSS sheared DNA library		
COMMENT	Unpublished (2000)		
	Contact: Brendan J Loftus		
	Department of Eukaryotic Genomics		
	The Institute for Genomic Research		
	9712 Medical Center Dr., Rockville, MD 20850, USA		
	Tel: 301 838 0208		
	Email: bjoftus@tigr.org		
	Fax: 301 838 3543		
	Clones are derived from the Entamoeba histolytica HM1:IMSS sheared DNA library		
	Seq primer: ML3-Forward		
	Class: shotgun		
	High quality sequence start: 16		
	High quality sequence stop: 860.		
FEATURES	Location/Qualifiers		
source	1..871		
	/organism="Entamoeba histolytica"		
	/strain="HM1:IMSS"		
	/db_xref="taxon:5759"		
	/clone_lib="Entamoeba histolytica Sheared DNA"		
	/note="Vector: pHOS1; Site 1: Bst I; Constructed at Th		

QY 1003 LeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSer 1022
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 Db 219 CTGGAGGTGAAGAAAGCCAGACACTAGCCTGGTGGGAGCAGTGGCTCTGGGAAGAGC 278
 QY 1023 ThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAsp 1042
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 Db 279 ACGGTGGTCCAGCTCTGGAGCGGTTCTACGACCCCTTGGCGGGACAGTGGCTTCTCGAT 338
 QY 1043 GlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaLeuValPro 1062
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 Db 339 GGTCTAGAGCAAGAAACTCAATGTCCAGTGGCTCAGAGCTCACTCGGAATCGGTCT 398
 QY 1063 GlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSer 1082
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 Db 399 CAGGAGCCTATCTATTGACTGCAGCATTTGCCGAGATATTGCTTATGGAGACAACAGC 458
 QY 1083 ArgValValProLeuAspGluIleLysGluAlaAsnAlaAsnIleHisSerPhe 1102
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 Db 459 CGGGTGTATCATCAGATGAATCTGTGAGTGCAGCCAAAGCTGCCAACATACATCTTTC 518
 QY 1103 IleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSer 1122
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 Db 519 ATCGAGACGTTACCCACAAATATGAACAGAGTGGGAGATAGGGGACTCAGCTCTCA 578
 QY 1123 GlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLys-ProlysisIleLe 1142
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 Db 579 GGAGGTTCACAAACAGAGGATTGCTATTGCCGAGCCTCATCAGACAACCTCAAAATCT 638
 QY 1142 uLeuLeuAspGluAlaThrSerAla-LeuAspAsnAspSerGluLysValVal-----G 1160
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 Db 639 CCTGTTGGATGAAGTACATCAGCTCTGATACCTGCAGCAAGTGAACACGGTTGTCCCA 698
 QY 1160 lnhIsAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuS 1180
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 Db 699 GAAGCCCTGTACACAGACAGAGCGCGACCTGCTTGTGATGTCTCACCCTGTC 758
 QY 1180 exAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLeuGluGlnG 1200
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 Db 759 AAC--ATCCAGATGCAGACTT-ATAGTGTGTGCCAACGGGAGAAGTCAAG----- 808
 QY 1200 lyThrHis 1202
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 Db 809 --ACACAT 814
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 RESULT 15
 AK020318
 LOCUS
 DEFINITION Mus musculus adult male epididymis cDNA, RIKEN full-length enriched library, clone:9230106F14:similar to MULTIDRUG RESISTANCE PROTEIN 3 (P-GLYCOPROTEIN 3) (MDR1A), full insert sequence.
 ACCESSION AK020318
 VERSION AK020318.1 GI:12860872
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (strain:C57BL/6J) adult male epididymis cDNA to mRNA, clone lib:RIKEN full-length enriched mouse cDNA library
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1
 AUTHORS Carninci, P. and Hayashizaki, Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636
 REFERENCE 2
 AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE 20499374

11042159

PUBMED
REFERENCE
AUTHORS

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuina, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Maeumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 20530913
 11076861

PUBMED
REFERENCE
AUTHORS

4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojibori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, H., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyo-oka, K., Wang, K. H., Weitz, C., Whittaker, C. C., Wilmfing, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.

Functional annotation of a full-length mouse cDNA collection
 Nature 409 (6821), 685-690 (2001)

PUBMED
REFERENCE
AUTHORS

5 (bases 1 to 872)
 21085660
 11217851
 Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hirooka, T., Hori, F., Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Schriml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamanaka, I., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

PUBMED
REFERENCE
AUTHORS

Direct Submission
 Submitted (18-AUG-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
 Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'-GACGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTT 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 185.2. Second strand cDNA was prepared with the primer

adapter of sequence [5'
CAGAGAGAGATCTCGAGTAAATTAATATCCGCCCCCCCCC 3']. cDNA was cleaved
with BamHI and XhoI. Vector: a modified pBluescript KS(+) after
bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3'
end: BamHI. Host: DH10B

FEATURES

source

Location/Qualifiers

1..872
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="FANTOM DB:9230106F14"
/db_xref="MGD:MGI:1911565"
/db_xref="taxon:10090"
/clone="9230106F14"
/sex="male"
/tissue type="epididymis"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"

CDS

86..>872
/notes="data source:SPTR, source key:P21447, evidence:ISS
putative
similar to MULTIDRUG RESISTANCE PROTEIN 3 (P-GLYCOPROTEIN
3) (MDR1A)"
/codon_start=1
/protein_id="BAB32070.1"
/db_xref="GI:12860873"
/db_xref="MGD:MGI:1924956"
/translation="MANERTNGLOETNORYGLOEVPKGVQNAVGPTEIFRFADNL
DIVLMTGLIASINGATVPLASVLGEISDHLINGLAVQTNRTKYQNCSTQEKLE
DIIVLYYIGIGAAALIFGYVOISFWITPAARQTRIRKQFPHSILAODISWFGSD
ICELNTRMTGDINKLGDGIDKPLMFQNIISGIFSLIGLISLKSWSKLSLVILSTPLI
MASSALCSRMIIISLTSEKLDAYSKAVAEALSSITQVTAFGA"

BASE COUNT 262 a 178 c 187 g 245 t
ORIGIN

Alignment Scores:

Pred. No.:	1.69e-70	Length:	872
Score:	744.50	Matches:	151
Percent Similarity:	80.37%	Conservative:	21
Best Local Similarity:	70.56%	Mismatches:	27
Query Match:	12.09%	Indels:	15
DB:	11	Gaps:	2

US-09-873-409-5 (1-1222) x AK020318 (1-872)

Qy	1	MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu	20
Db	230	ATGACTCTGGGAATATTAGCATTCATGATAAATGAGCCCGCTCCTTTAATGTCCTG	289
Qy	21	ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr	40
Db	290	GTTTTAGGAGAAATAAGTGATCATTTAATTAAATGATGCCTAGTACAAACTAACAGAACT	349
Qy	41	SerPhe-----PheArgLeuThr	46
Db	350	AAATATCAGAACTGTTCTCAGACTCAAGAAAGCTGAATGAAGATATCATTTGTTGACT	409
Qy	47	LeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeu	66
Db	410	CTATATTATATTGGAATAGGAGCAGCTGCCCTCATTTTGGCTATGTACAGATTTCCTC	469
Qy	67	TrpIleIleThrAlaAlaArgGlnThrIysArgIleArgLysGlnPhePheHisSerVal	86
Db	470	TGGTCATACTGCAGCCCGCAACCAACAGAGATCCGMAAACAGTTTTTTCATTCATT	529
Qy	87	LeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGluLeuAsnThrArgMet	106
Db	530	TTGGCACAAGACATCAGCTGGTTTGTATGGCAGTGCATCTGTGAACCTTAACACCCGCATG	589
Qy	107	Thr---AspIleAspIleIleSerAspGlyIleGlyAspIleIleAlaLeuLeuPheGln	125
Db	590	ACTGGTGACATCAACAAACTCTGTGATGGTATTTAGATAAGATCCCTCTGATGTTTCAG	649
Qy	126	AsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLysGlyTyrLysLeuThr	145

Db	650	AAACATATCTGGGTTTTCTATTGGCTTGGTATAAGTTTATATAAAAGCTGGAAACTCTCC	709
Qy	146	LeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAlaCysSerArgMet	165
Db	710	CTGGTGGTTCTGTCTACATCTCTCTCATATGGCTTCATCGGCACCTGTGCTTAGGATG	769
Qy	166	ValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGlu	185
Db	770	ATTATCTCATTCAGCAGCAAGAGCTGGATGCTATTTCAAAGCTGGGGCTGTGGCTGAA	829
Qy	186	GluValLeuSerSerIleArgThrValIleAlaPheArgAla	199
Db	830	GAAGCCTTGTCTCAATCAATCAAAACAGTCACAGCCTTTGGAGCC	871

Search completed: March 31, 2003, 13:50:58
Job time : 8432.26 secs

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 30, 2003, 01:20:01 ; Search time 1194.3 Seconds
(without alignments)
2304.231 Million cell updates/sec

Title: US-09-873-409-5

Perfect score: 6159

Sequence: 1 MITGLIASLVNGACPLMLP.....QSLRLNRDIYFKLVNAQSVQ 1222

Scoring table:
BLOSUMP62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-O=cpn2_1/USPTO spool/US09873409/runat_27032003_115416_19216/app_query.fasta_1.7544
-DB=N_Geneseq_101002 -QPM=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09873409 @CNC 1 1 4247 @runat_27032003_115416_19216 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DSV TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_101002.*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Length	DB	ID	Description
1	5737	93.1	3699	24	ABK83223	Human transporter
2	3533.5	57.4	4279	22	AA03506	Dog P-glycoprotein
3	3533	57.4	4195	22	AAF86128	Cynomologous monke
4	3530.5	57.3	3860	21	AAZ49332	Human wild-type mu
5	3530.5	57.3	3860	24	ABA94365	Human BCRP DNA rel
6	3530.5	57.3	4279	22	AA03504	Dog P-glycoprotein
7	3530.5	57.3	4279	22	AA03505	Dog P-glycoprotein
8	3525.5	57.2	4279	22	AA03488	Dog P-glycoprotein
9	3525.5	57.2	4317	22	AA03489	Dog P-glycoprotein
10	3524.5	57.2	4186	22	AAF86127	Cynomologous monke
11	3521.5	57.2	3860	21	AAZ49333	Human G185V mutant
12	3521.5	57.2	3860	24	ABA94366	Human BCRP DNA rel
13	3521.5	57.2	4349	22	AAH57442	Human intestine ce
14	3521.5	57.2	4646	21	AAZ94738	Human ATP binding
15	3521.5	57.2	4646	24	AAZ38994	Human mdrl gene.
16	3521.5	57.2	4646	24	ABL68592	Kidney cancer rela
17	3521.5	57.2	4646	24	ABL68880	Kidney cancer rela
18	3521.5	57.2	6505	17	AAAT13394	Hybrid vector pSF-
19	3521.5	57.2	8630	21	AAZ24041	Retroviral M4 mdr-
20	3521.5	57.2	8630	21	AAZ24042	Hybrid vector pSF-
21	3520	57.2	4264	19	AAV66533	Mutated human P-g
22	3520	57.2	4264	19	AAV66534	Mutated human P-g
23	3519.5	57.1	4669	8	AAV70752	Sequence of human
24	3517	57.1	4425	21	AAZ52048	Rat multidrug resi
25	3515.5	57.1	4378	11	AAQ04522	Multidrug Resistan
26	3515	57.1	4369	21	AAZ52047	Rat multidrug resi
27	3514	57.1	3988	21	AAZ89167	Human polynucleoti
28	3514	57.1	3840	24	ABL91687	Human MDR-1 DNA.
29	3512.5	57.0	4669	14	AAQ52726	Sequence of human
30	3510.5	57.0	4646	15	AAQ72872	Human multidrug re
31	3503.5	56.9	4669	19	AAV32645	Human P glycoprote
32	3503.5	56.9	4669	21	AAK52041	cDNA encoding huma
33	3503.5	56.9	4189	21	AAZ49334	Murine multidrug r
34	3503	56.9	4189	24	ABA94367	Mouse BCRP DNA rel
35	3503	56.9	4313	14	AAQ38950	Mouse multidrug re
36	3499	56.8	4788	21	AAZ49335	Murine multidrug r
37	3499	56.8	4788	24	ABA94368	Mouse BCRP DNA rel
38	3494	56.7	4233	21	AAZ90198	Rat mdrlb2 (multis
39	3494	56.7	4233	22	AAF27498	Rat mdrlb2 multidr
40	3481	56.5	3924	21	AAZ94742	Human ATP binding
41	3481	56.5	3924	21	AAZ88974	Human MDR-3 DNA.
42	3481	56.5	3924	24	ABN95801	Gene #2299 used to
43	3453	56.1	4254	24	ABK63517	Rat sequence diffe
44	3405	55.3	3912	24	ABK63653	Rat sequence diffe
45	2844.5	46.2	4776	21	AAZ94744	Human ATP binding

ALIGNMENTS

RESULT 1	
ABK83223	
ID	ABK83223 standard; cDNA; 3699 BP.
XX	ABK83223;
AC	
XX	
XX	27-AUG-2002 (first entry)
DT	
DE	Human transporter and ion channel, TRICH14, Incyte ID 7472030CB1, cDNA.
XX	
KW	Human; ss; gene; transporter and ion channel; TRICH; transport disorder;
KW	neurological disorder; muscle disorder; immunological disorder; cancer;
KW	scleroderma; systemic lupus erythematosus; allergy; leukaemia;
KW	cell proliferative disorder; cervical cancer; breast cancer;
KW	neurodegenerative disorder; Parkinson's disease; Alzheimer's disease;
KW	myoclonic dystrophy; catatonia; endocrine disorder; diabetes;
KW	Grave's disease; gastrointestinal disorder; Crohn's disease;
KW	renal disorder; Good pasture's syndrome; viral infection; cirrhosis;

KW bacterial infection; fungal infection; parasitic infection;
 KW protozoal infection; helminthic infection; cardiovascular disorder;
 KW atherosclerosis; hepatic disease.

XX Homo sapiens.

XX WO200240541-A2.

XX 23-MAY-2002.

XX 25-OCT-2001; 2001WO-US46055.

XX 27-OCT-2000; 2000US-243989P.

PR 03-NOV-2000; 2000US-245904P.

PR 09-NOV-2000; 2000US-247673P.

PR 17-NOV-2000; 2000US-249661P.

PR 20-NOV-2000; 2000US-252232P.

PR 01-DEC-2000; 2000US-250790P.

XX (INCY-) INCYTE GENOMICS INC.

XX Tang YT, Yue H, Nguyen DB, Hafalia AJA, Elliott VS, Lu Y;

PI Wallia NK, Yao MG, Baughn MR, Gandhi AR, Ding L, Sanjanwala M;

PI Ramkumar J, Arvizu C, Gietzen KJ, Lal PG, Azimzai Y, Khan FA;

PI Thangavelu K, Thornton M, Lu DAM, Tribouley CM, Warren BA;

PI Ison CH, Das D, Raumann BE, Policky JL, Kearney L;

XX WPI; 2002-463570/49.

DR P-PSDB; ABG61544.

XX New transporters and ion channels (TRICH) polypeptides, useful for
 PT diagnosing, preventing, and treating disorders associated with an
 PT abnormal expression or activity of TRICH, e.g. immunological, muscular
 PT or renal disorders

XX Claim 5; Page 172; 179pp; English.

XX The invention relates to human transporters and ion channels (TRICH)
 CC polypeptides, a naturally occurring amino acid sequence 90 % identical to
 CC TRICH, a biologically active fragment of TRICH or an immunogenic fragment
 CC of TRICH. Also included are an isolated polynucleotide encoding TRICH,
 CC a recombinant polynucleotide comprising a promoter sequence operably
 CC linked to the TRICH polynucleotide, a cell transformed with the
 CC recombinant polynucleotide, a transgenic organism comprising the
 CC recombinant polynucleotide, an isolated antibody that binds specifically
 CC to TRICH, and screening for compounds which bind to TRICH, modulate
 CC TRICH, modulate TRICH expression or are ant/agonists of TRICH.
 CC The polypeptides are useful for diagnosing, treating, and
 CC preventing transport, neurological, muscle, immunological disorders
 CC (e.g. scleroderma, systemic lupus erythematosus, allergies), cell
 CC proliferative disorders such as cancers (e.g. leukaemia, cervical or
 CC breast cancers), neurodegenerative disorders (e.g. Parkinson's disease,
 CC Alzheimer's disease), muscular disorders (e.g. myotonic dystrophy,
 CC catatonla), endocrine disorders (e.g. diabetes, Grave's disease),
 CC gastrointestinal disorders (e.g. Crohn's disease), renal disorders
 CC (e.g. Good pasture's syndrome), vital, bacterial, fungal, parasitic,
 CC protozoal and helminthic infections, cardiovascular disorders (e.g.
 CC atherosclerosis), or hepatic diseases (e.g. cirrhosis) and many
 CC other diseases and disorders detailed in the specification. They can also
 CC be used in assessing the effects of exogenous compounds on the
 CC expression of nucleic acid and amino acid sequences of transporters and
 CC ion channels. TRICH or its fragments may also be used in screening for
 CC compounds that specifically bind to and modulate the activity of TRICH.
 CC The polynucleotides can be used to create knock-in humanised animals or
 CC transgenic animals to model human disease. The present sequence
 CC encodes a TRICH protein.

XX Sequence 3699 BP; 1116 A; 707 C; 860 G; 1016 T; 0 other;

Alignment Scores:

Prd. No.: 0 Length: 3699
 Score: 5737.00 Matches: 1156
 Percent Similarity: 93.78% Conservative: 4

Best Local Similarity: 93.45% Mismatches: 9
 Query Match: 93.15% Indels: 68
 DB: 24 Gaps: 4

US-09-873-409-5 (1-1222) x ABK83223 (1-3699)

QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
 Db 145 ATGATCTGGGTATACCTGACATCAGTCTTCAATGGAGCCTGCTCTTAAATGCCACTG 204
 QY 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
 Db 205 TGTATAGGAGAAATAGTATACCTTATTAGTGGATGTCTAGTCCACACTAAACACCA 264
 QY 41 SerPhePheArg-----LeuThr 46
 Db 265 AATTATCAGAACTGACTCAGTCTCAAGAGAGCTCAATGAAGATATGACTCTGTTGACC 324
 QY 47 LeuTyrTyrValGlyIleGlyValAlaLeuIlePheGlyTyrIleGlnIleSerLeu 66
 Db 325 CTGTATTATGTTGGAATAGTGTGCTGCTTGAATTTTGGTTACATACAGATTCTCTTG 384
 QY 67 TrpIleIleThrAlaAlaArgGlnThrLysArgIleAArgLysGlnPhePheHisSerVal 86
 Db 385 TGGATTATTAACTCGACGACGACAGACCAAGAGAGATTTCGAAAAACAGTTTTTTCATT 444
 QY 87 LeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGluLeuAsnThrArgMet 106
 Db 445 TTGGCACAGACATCGGCTGGTTTGTATAGTCTGACATCGGTGAACCTTAACATCCGATG 504
 QY 107 Thr---AspIleAspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGln 125
 Db 505 ACAGATGACATTGACAAAATCAGTGATGTTATGGAGATAAGATTCTCTGTGTTTCAA 564
 QY 126 AsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLysGlyTrpLysLeuThr 145
 Db 565 AACATCTCTACTTTTTTCGATTGGCTGGCAGTTGGTGTGTAAGGGCTGGAAACCTCACC 624
 QY 146 LeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAlaAlaCysSerArgMet 165
 Db 625 CTAGTGACTCTATCCACGTCCTCTTATTAAGTGGCTTCAGCGGCAGCATGTTCTAGATG 684
 QY 166 ValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGlu 185
 Db 685 GTCATCTCTATTCACAGTAGTAAGGAATTAAGTGCCTATTCAAAGCTGGGCTGGCAGAA 744
 QY 186 GluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGlyLysGluLeuGln 205
 Db 745 GAAGTCTTGTTCATCAATCCGAACAGTCATAGCCTTTAGGGCCCGCAGGAGAAAGAACTCAA 804
 QY 206 ArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeu 225
 Db 804 -----AGGTATACACAGAACTCAAGATGCAAAAGGATTGTTGGC 843
 QY 226 SerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 245
 Db 805 -----AGGTATACACAGAACTCAAGATGCAAAAGGATTGTTGGC 843
 QY 246 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGly 265
 Db 844 ATAAAAAGGACTATAGCTCTCAAAAGTGTCTCTTGGTGGCTGTGTACTCTTTATGAATGA 903
 QY 266 ThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 285
 Db 904 ACCTATGGACTTGTCTTTTGGTATGGAACCTCTTGTATCTTATATGAGAACCTGGATAT 963
 QY 286 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 305
 Db 964 ACCATCGGACTGTCTTCTGCTGTTTCTTTAGTGTAAATCCATAGTAGTTATTTCATTGGA 1023
 QY 306 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 325
 Db 1024 GCAGCAGTCCCTCCTTGAACCTTTCGCAATAGCCCGAGGAGCTGCTTTCATATTTTC 1083

326 GlnValIleAspLysPheSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 345
 1084 CAGGTTATTGATAAGAAACCAGTATAGGTAACTTTTCCACAGCTGATATAAACCCTGAA 1143
 346 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 365
 1144 TCCATAGAAGAACTGTGGAATTTAAATAATGTTCTTTCAATATCCATCAAGGCCATCT 1203
 366 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 385
 1204 ATCAAGATTCTGAAGGTCGTGAATCTCGGAATTAAGTCTGGAGAGACAGTCGCTTGGTC 1263
 386 GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro 405
 1264 GGTCTCAATGGCAGTGGGAAGAGTACGTAGTCCAGCTTCTGCAGAGGTATATGATCCG 1323
 406 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 425
 1324 GATGATGGCTTTATCATGTGGTGGATGGAATGACATCAGAGCTTTAAATGTGCGGCATTAT 1383
 426 ArgAspHisIleGlyValValSerGlnProValLeuPheGlyThrThrIleSerAsn 445
 1384 CGAGACCATATTGGAGTGGTGTAGTCAAGAGCCTGTTTGTTCGGACCCACCATCAGTAAC 1443
 446 AsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGlu 465
 1444 AATATCAAGTATGACGAGATGATGTGACTGATGAAGAGATGAGAGAGCAGCAAGGAA 1503
 466 AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 485
 1504 GCAAAATCGGTATGATTTTATCATGGAGTTCCTTAATAAATTAATCATTTGGTAGGGAA 1563
 486 LysGlyValAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal 505
 1564 AAGAGAGCTCAATGATGAGTGGAGGCGAGAAACAGAGGATCCCAATGTCTGTGCTTAGTT 1623
 506 ArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 525
 1624 CGAAACCCCAAGATCTGATTTTAGATGAGGCTACGCTCGCCCTGGATTCAAGAAAGCAAG 1683
 526 SerAlaValGlnAlaLeuLeuLysAlaSerLysGlyArgThrThrIleValValAla 545
 1684 TCAGCTGTTTCAAGCTGACCTGGAGAGCGCAGCAAGCTCGGACTACAATCGTGGTAGCA 1743
 546 HisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeu 565
 1744 CACCGACTTCTACTATTCGAAGTGCAGATTGTAATGTGACCCCTAAAGGATGGAATGCTG 1803
 566 AlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuVal 585
 1804 GCGAGAAAGAGGCGCATGCTGAACCTAATGCGAAACAGAGTCTATATTATTCACCTGTG 1863
 586 MetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu 605
 1864 ATGTCAAGGATATTAATAAAGCTGTATGAACAGAGTGAATGAGTCAATGACATATCTACTGAA 1923
 606 ArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIleAsp 625
 1924 AGAAGACCAACTCATTCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTCATTGAC 1983
 626 LysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLysLys 645
 1984 AAGCTGAGGAAATCCACCAATCTAAGAGATAGTCTTCTGAGTCTCTCTATTANAA 2043
 646 IleLeuLysLeuAsnLysProGluTyrProPheValValLeuGlyThrLeuAlaSerVal 665
 2044 ATTTAAAGTTAAACAAGCTGAATGCCCTTTGTGGTCTCTGGGACATTTGGCTTCTGTT 2103
 666 LeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPhe 685
 2104 CTAATGGAACTGTTCATCCAGTATTTTCCATCATCTTTGCAAAAAATTAACCATGTTT 2163

686 GlyAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheVal 705
 2164 GGAATAATATGATAAAACCAACCATTAAGCATGATGCAGAAATTTATTTCCATGATATTCGTC 2223
 706 IleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAla 725
 2224 ATTTTGGGTGTTATTTGCTTTGTTCAGTTATTTTCATG----- 2259
 726 GlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAsp 745
 2260 -----CAGGAT 2265
 746 IleAlaTyrPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIle 765
 2266 ATTCCTGCTGTTTGTGATAAAGGAAACACACACAGGAGGCTTGACAAATATTTAGGCATA 2325
 766 AspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAla 785
 2326 GATATAGCAAAATTCAGGAGCAACAGGTTCCAGATTGGCGTCTTTAACACAAAATGCA 2385
 786 ThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyrGluMetThrPheLeu 805
 2386 ACTAAACATGGACTTTTCAGTTATCATTTCTTATATATGATGGAGATGACATTTCTCTG 2445
 806 IleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThr 825
 2446 ATTTCTGATTGCTCCAGTACTTGCCTGACAGGAATGATTGAAACCGCAGCAATGACT 2505
 826 GlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAla 845
 2506 GGATTTGCCAAACAAAGATAGCAAGAACTTTAAGCATGCTGGAAAGATAGCAACTGAAGCT 2565
 846 LeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyr 865
 2566 TTGGAGAATATACGTACTATAGTGTCAATTACACAGGAGAAAGCCTTCGAGCAATGTAT 2625
 866 GluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySer 885
 2626 GAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCACAGATTATTGGAAGC 2685
 886 CysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGly 905
 2686 TGTATGCAATTCAGCATGCTTATATATTTTGCCTATGCGCAGGGTTTCGATTGGA 2745
 906 AlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAla 925
 2746 GCCTATTTAATTCAGCTGGACGAATGACCCAGAGGGCATGTTTCATAGTTTACTGCA 2805
 926 IleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLys 945
 2806 ATTCATATGGAGCTATGGCCATCGGAGAAACGCTCGTTTGGCTCTCGTAATATTCAAA 2865
 946 AlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSer 965
 2866 GCCAAATCGGGGCTCGCATCTGTTTGCCTTGTTCGAAAAGAAACCAATATAGACAGC 2925
 966 ArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluVal 985
 2926 CGCAGTCAAGAGGAGAAAGCCAGACACATGTGAAGGGAATTTAGAGTTTCGAGAGTC 2985
 986 SerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIle 1005
 2986 TCITTTCTCTCATCATGTCGCCAGATGTTTTTCATCTCGTGGCTTATCCCTCAGTATT 3045
 1006 GluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerVal 1025
 3046 GAGCGAGGAAGACAGTAGCATTTGTGGGAGCAGCGGCTGTGGGAAAGACACTTCTGTT 3105
 1026 GlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAsp 1045
 3106 CAACCTTCGAGAGACTTTATGACCCCGTCGCAAGGACAGTGTCTGTTGATGTGTGGAT 3165
 1046 AlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluPro 1065


```
Db 3166 GCAAAAGATTGAATGTACAGTGGCTCGGTTCCAAATAGCAATCGTTCCTCAAGAGCCT 3225
Qy 1066 ValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValVal 1085
Db 3226 GTGCTCTTCAACTGCACGATGTGTGAGAACATCGCTTATGGTACACACCGCTGTGGTG 3285
Qy 1086 ProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGly 1105
Db 3286 CCATTAGATGAGATCAAGAAGCGCAATGCGCAATATCCATCTTTATTATGAAGGT 3345
Qy 1106 LeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGln 1125
Db 3346 CTCCTTGAGAAATACACACACAGATTTGGACTGAAAGGAGCACAGCTTTCTGCGGCCAG 3405
Qy 1126 LysGlnArgLeuAlaIleAlaArgAlaLeuGlnLysProLysIleLeuLeuLeuAsp 1145
Db 3406 AACAAGAGCTAGCTATGTGAAGGGCTCTTCTCCAAAACCCAAAATTTATTATTGGAT 3465
Qy 1146 GluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLys 1165
Db 3466 GAGGCCACTTCAGCCCTCGATATGACAGTGAGAGGTGGTTGAGCATGCCCTTGATAAA 3525
Qy 1166 AlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAla 1185
Db 3526 GCCAGGACGGAGGACATGCTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAACGCA 3585
Qy 1186 AspleuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeu 1205
Db 3586 GATTGTGATGTGGTTCTGCAATGGAAGATTAAGGAACAGCACTCATCAAGAGCTC 3645
Qy 1206 LeuArgAsnArgAspIleTyrPheLeuLeuValAsnAlaGlnSerValGln 1222
Db 3646 CTGAGAAATCGAGACATATATTTTAAGTTAGTGAATGCACAGTCAGTGCAG 3696

RESULT 2
AAD03506
ID AAD03506 standard; cDNA; 4279 BP.
XX
AC AAD03506;
XX
DT 13-JUN-2001 (first entry)
DE
DE Dog P-glycoprotein (PGP) allelic variant (Genotype D) cDNA.
XX
XX Dog; P-glycoprotein allelic variant; PGP; multidrug transporter;
KW MDRI; drug bioavailability; transgenic animal; genetic model; ss.
XX
OS Canis familiaris.
XX
FH Key Location/Qualifiers
FT CDS 17..3862
FT /*tag= a
FT /product= "Dog P-glycoprotein (PGP) allelic variant
FT (Genotype D) protein"
FT allele replace (91, T)
FT /*tag= b
FT allele replace (607, C)
FT /*tag= c
FT allele replace (1001, T)
FT /*tag= c
FT allele replace (3459, A)
FT /*tag= c
XX
XX WO200123540-A2.
XX
XX 05-APR-2001.
XX
XX 28-SEP-2000; 2000WO-US26767.
XX
XX 28-SEP-1999; 99US-0156510.
XX
XX (GENT-) GENTEST CORP.
```

```
XX
PI Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;
XX WPI; 2001-235373/24.
DR P-PSDB; AAE00310.
XX
PT New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
PT for determining the bioavailability of drugs and for screening for dog
PT PGP inhibitors -
XX
PS Claim 9; Page 102-107; 11pp; English.
XX
CC The invention relates to dog P-glycoprotein (PGP) also referred
CC as multidrug transporter (MDRI) and nucleic acids encoding them.
CC The invention also includes fragments and biologically functional
CC variants of dog P-glycoprotein. PGP and their nucleic acids are
CC useful for determining the bioavailability of drugs and for
CC screening PGP inhibitors. They are useful for the diagnosis and
CC treatment of conditions characterised by PGP activity, by
CC reducing or increasing PGP activity in a cell. PGP nucleic acids
CC are used as oligonucleotide probes. Complements of PGP nucleic
CC acids are useful as antisense oligonucleotides, to induce a PGP
CC 'knockout' phenotype. They are used to prepare a non-human
CC transgenic animal, which are valuable as genetic models for
CC human diseases.
CC The present sequence is dog P-glycoprotein (PGP) allelic variant
CC (Genotype D) cDNA. The PGP enzyme functions as an efflux pump
CC exporting small molecules across the cell membrane. This enzyme
CC is a member of the ABC transporter family.
XX
SQ Sequence 4279 BP; 1296 A; 833 C; 1009 G; 1141 T; 0 other;

Alignment Scores:
Pred. No.: 0 Length: 4279
Score: 3533.50 Matches: 683
Percent Similarity: 73.13% Conservative: 237
Best Local Similarity: 54.29% Mismatches: 269
Query Match: 57.37% Indels: 69
DB: 22 Gaps: 9

US-09-873-409-5 (1-1222) x AAD03506 (1-4279)
Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 167 ATGTTGGTGGGACAAATGCTGCCATCATCGAGTGCACCTCCCTCTCATGATGCTG 226
Qy 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu----- 34
Db 227 GTTTTGGAAACATGACAGATAGCTTTGCCAAATGCGAGGAATTTCAAGAAACAAACTTTT 286
Qy 35 -----ValGlnThrAsnThrTyrSerPhePheArg----- 44
Db 287 CAGTTATTAATTAATGAAAGTATTACGAACATATACAACTTTTCATCAACCATCTGGAG 346
Qy 45 -----LeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePhe 59
Db 347 GAGGAATGACCAAGGATGCTTATTATTACAGTGGGATCGGTGCTGCGTGTGGTGGCT 406
Qy 60 GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg 79
Db 407 GCTTACATCCAGGTTTCATTTCTGCTGCGCAGGAGGAGACAGATACTCAAAATTAGA 466
Qy 80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99
Db 467 AACAAATTTTTCATGCTATCATGCGACAGGAGATTTGGCTGGTTTACCGTGCATGACGTT 526
Qy 100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
Db 527 GGGAGAGCTTAACACCCGGCTCACAGACGATGTCTCCAAATCAATGAGGAATTGGGAC 586
Qy 119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
Db 587 AAAATTGGAATGTTCTTCTTCAATCAATGAGCAACATTTTTCACCGGTTTATATAGTGGGTTT 646
```


QY 139 ValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuLeuMetAlaSer 158
Db 647 ACACGTTGGTGAAGCTAACCCCTGTGATTTTGGCCATCAGCCCTGTCTTGGACATTCA 706
QY 159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
Db 707 GCGCCATCTGGCCAAAGATACCTTCTTACTGATAAAGAACTCTTGGCCTATGCA 766
QY 179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198
Db 767 AAAGCTGGAGCAGTACTGAGAAAGCTTACAGCAATCAGAACTGTGATTGCTTTGGA 826
QY 199 AlaGlnGlyLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPhe 218
Db 827 GGCAAAAGAAAGAACTTGAA----- 847
QY 219 TyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeu 238
Db 848 -----AGGTACAACAAAAATTTA 865
QY 239 LysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAla 258
Db 866 GAAGAAGCTTAAGGAATTGGGATTAAGAAAGCTATCAGGCCCAACATTTCTATTGTGCC 925
QY 259 ValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIle 278
Db 926 GCATTCTTATGATCATGCATCATATGCTCTGCCTTCTGGTATGGACCTCTCTGGTC 985
QY 279 LeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIle 298
Db 986 CTCTCCAGTGAA-----TATACTATTGGACAAGTACTCACTGCTCTCTTTCTGTATTA 1039
QY 299 HisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArg 318
Db 1040 ATGGGGCTTTTATGATTGGACAGGATCCCAAGCAATTGAAGCATTTGCAAAACGCAAGA 1099
QY 319 GlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSer 338
Db 1100 GGAGCAGCTTATGAATCTTCAGAGATAATTGACATAAACCAAGCATTCACACGTATTCG 1159
QY 339 ThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPhe 358
Db 1160 AAGAGTGGACATAAAACAGATAATATTAAAGGAAATTTGGAATTTCAAAAATGTTCACTTC 1219
QY 359 AsnTyrProSerArgProSerIleLysIleLysGlyLeuAsnLeuArgIleLysSer 378
Db 1220 AGTTACCCCTTCTCGAAAGAAAGTAAAGATCTTAAAGGGTCTCAACCTGGAAGGTTACAGT 1279
QY 379 GlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeu 398
Db 1280 GGGCAGACAGTGGCGCTGTTGGACAGTGGCTGGCGGAAGACGACGCGGTGCAGCTG 1339
QY 399 LeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArg 418
Db 1340 ATGCAGAGGCTCTATGACCCACAGATGGCATGCTCTGTATTATGATGACAGGACATTAGG 1399
QY 419 AlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeu 438
Db 1400 ACCATAAATGTAAGGCATCTTCGGGAAATTTACTGGTGTGTGAGTACGAGGCTGTGTTG 1459
QY 439 PheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGlu 458
Db 1460 TTTGCCACACGATAGCTGAAACATTCGCTATGGCCGCCGCAAAATGTCACCATGGATGAG 1519
QY 459 MetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLys 478
Db 1520 ATTGAGAAAGCTGTTAAGGAAGCAATGCCTATGATTTTATGAAACTACCTAATAAA 1579
QY 479 PheAsnThrLeuValGlyGlyLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIle 498
Db 1580 TTTGACACTCTGTTGGAGAGAGGGGCCCGCAGCTGAGTGGTGACAGAAACAGAGAAATC 1639

QY 499 AlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSer 518
Db 1640 GCCATTGCTCGGCCCTGTTCCCAACCCCAAGATTCTTCTGCTGATGAGGCAACGTCA 1699
QY 519 AlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGly 538
Db 1700 GCTCTGGACACTGAAAGTGAAGCAGTGGTTCAGGTGCCCTCGATGAAGGCCAAGAAAGGC 1759
QY 539 ArgThrThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleVal 558
Db 1760 CGGACTACCATTTGTGATAGCTCATCGTTTGTCTACAGTTCTGTAATGCCGATGTCATTGCT 1819
QY 559 ThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArg 578
Db 1820 GGTTTTGTATGATGGATCATCTGTGGAGAAAGGAAATCATGTAACATCATGAAGAAGAG 1879
QY 579 GlyLeuTyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGlu 598
Db 1880 GGCATTACTTCAAACTTGTCAATGCAG---ACNAGAGGAAATGAAATTTAGTTAGAA 1936
QY 599 SerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSer 618
Db 1937 AATGCCACTGGTGAATCCAAAGTGAAGTGCCTTGGAAATG---TCTCCAAAAGAT 1993
QY 619 IleLysSerAspPheIle----- 624
Db 1994 TCAGGCTCCAGTTTAAATAAAAGAGATCAACTCGCAGGAGTATACATGCACCACCAAGGC 2053
QY 625 ---AspLysAlaGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeu 643
Db 2054 CAAGCAGAAAGCTGGTGTACAAAGAGAGCTTGAATGAGAATGTACTCCAGTTTCCCTTC 2113
QY 644 LeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAla 663
Db 2114 TGGAGATTTCTAGCTGAATCACTCACTGAATGGCCTATTATTGTGGTGGTATTTTGT 2173
QY 664 SerValLeuAsnGlyThrValHisProValPheSerIlePheAlaLysIleIleThr 683
Db 2174 GCTATTATAAACCGAGCGCTGCAACCCAGCATTTTCAATAATATTTTCAAGGATATAGGG 2233
QY 684 MetPhe---GlyAsnAsnAspLysThrLeuLysHisAspAlaGluIleTyrSerMet 702
Db 2234 ATCTTTACCCGAGATGAGGATCTTGAAACCAACACAGCAAGTAGTAACATGTTTCTGTA 2293
QY 703 IlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyr 722
Db 2294 TGTGTTCTAGTCTCTGGAATATTCTTTTATTACATTTTCTCCAGGCTTCACATTT 2353
QY 723 GlyArgAlaGlyIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeu 742
Db 2354 GGCAAGCTGGGAGATCCTCACTAAGCGCTTCGATACATGGTTTTTTCAGATCCATGCTG 2413
QY 743 TyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrIle 762
Db 2414 AGACAGATGTGAGTGGTGTGATGACCCCTAAACACCACTGGAGCATTTGACACACAGG 2473
QY 763 LeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThr 782
Db 2474 CTTGCCAATGATCGCGCTCAAGTTAAAGGGCTATAGTTCCAGGCTGCTGTCATTACC 2533
QY 783 GlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMet 802
Db 2534 CAGAAATATAGCAATCTTGGACAGGCAATTTATATCTTAACTATGTTGTTGGCAATTA 2593
QY 803 ThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAla 822
Db 2594 ACATTTTACTCTTAGCAATTTGACCAATTTGACCAATTTGCAATAGCAGGAGTTGTTGAAATGAAA 2653
QY 823 AlaMetThrGlyPheAlaAsnLysAspLysGlnLeuLysHisAlaGlyLysIleAla 842
Db 2654 ATGTTCTGTGGCAAGCACTGAAAGATAGAAAGAGCTGAGAGAGCTGGGAAGATTGCT 2713
QY 843 ThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGlu 862

US-09-873-409-5 (1-1222) x AAF86128 (1-4195)

Qy	1	MetIleuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProIleuMetProLeu	20
Db	250	ATGTTGGTGGAACTTTTGGCTGCATCATCCATGGAGCTGGACTTCTCTCATGATGCTG	309
Qy	21	ValLeuGlyGluMetSerAspAsnLeuIleSer-----GlyCys	33
Db	310	GTGTTTGGAGCATGACGGATACCTTTGCMAATTCAGGAAATTTAGGAGATTTAGGAGCT	369
Qy	34	LeuValGlnThrAsnThrTyrSer-----	41
Db	370	CTGTTGTTTAAACAACACTAATACGACGTAATATACACTGATACAGTCCCGTCATGAATCTG	429
Qy	42	-----PhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIle	58
Db	430	GAGGAAGATATGACACGAGTATGCTATTATATACAGTGGAAATTTGGTGGGGTGCTGGT	489
Qy	59	PheGlyTyrIleGlnIleSerLeuTyrIleIleThrAlaAlaArgGlnThrLysArgIle	78
Db	490	CTGCTTACATTCAGGTTTCATTTTGGTCCTGGCAGCTGGGAAGACAATACACAAAATT	549
Qy	79	ArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAsp	98
Db	550	AGAAAACAGTTTTTTCATGCTATAATCGCACAGGAGATAGGCTGGTTGATGTGCACGAT	609
Qy	99	IleGlyGlnLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGly	117
Db	610	GTTGGGAGCTTAACACCCCGCTTCACAGATGATGCTCCCAAGATTAATGAAGGAATTCGT	669
Qy	118	AspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGly	137
Db	670	GACAAAATCGAATGTTCTTTCACTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGA	729
Qy	138	LeuValLysGlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAla	157
Db	730	TTTACACGCTGTTGGAAGCTAAACCTTGTGATTTTGGCCATCAGTCTCTTCTTGGACTG	789
Qy	158	SerAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyr	177
Db	790	TCAGCTGCAGCTCTGGCGAAGATACCTCTCTCATTTACTGATAAAGAACTCTTAGCTTAT	849
Qy	178	SerLysAlaGlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPhe	197
Db	850	GCATAAGCTGAGCAGTAGCTGGAAGAGTCTTGGCAGCAATTAGAACTGTGATTCGATTT	909
Qy	198	ArgAlaGlnLysGluLeuAlaArgSerPheLeuLeuAsnIleThrArgTyrAlaTyr	217
Db	910	GGAGGACAAAAGAAAGAACTCGAA-----	933
Qy	218	PheTyrPheProGlnTrpLeuLeuSerCysValLeu**PheValArgTyrThrGlnAsn	237
Db	934	-----AGGTACAACAAAAAT	948
Qy	238	LeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGly	257
Db	949	TTAGAAGACCTAAGAATTGGGATTAAGAAGACTATTACAGCCAAATTTTCTATAGGT	1008
Qy	258	AlaValTyrPhePheMetAsnGlyTyrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeu	277
Db	1009	GCTGCTTCTCGCTTATCTATGATCTTATGCTCTGGCTCTTGGTATGGGACCACTTG	1068
Qy	278	IleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerVal	297
Db	1069	GTCTCTTCAAGGAA-----TATTCTATGGCAAGTACTCACGTGATCTCTTTCTCTGA	1122
Qy	298	IleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAla	317
Db	1123	TTAATTGGGCTTTAGTGTGTGACAGGCATCTCCAAGCAATTGAACATTTGCAATGCA	1182
Qy	318	ArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPhe	337

Db	1183	AGAGGACGACGCTTTTGGAAATCTTCAAGATAAATTGATAATAAGCCAAAGTATTGACAGCTAT	124
Qy	338	SerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSer	357
Db	1243	TCGAAGAGTGGGCACAAACCAGATAATAATTAAGGGAAATTTGGAATTCAGAAATGTTCCAC	1302
Qy	358	PheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLys	377
Db	1303	TTCAAGTTACCCCATCTCGAAAGAAGTTAAGATCTTTGAAGGGCTCGAACTGAAGGTGAC	1362
Qy	378	SerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGln	397
Db	1363	AGTGGCAGACGGTGGCCCTGGTTGGAAACACGGCTGTGGGAAGACACAAACGGTCCG	1422
Qy	398	LeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIle	417
Db	1423	CTGATGCAGAGGCTTTATGACCCACAGAGGCGATGTCAGTGGTTGATGGACAGGATATT	1482
Qy	418	ArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProVal	437
Db	1483	AGGACCAATAACCTAAGGTTTCTACCGGGAATCATCGGTGTGTGAGTCAGGAACCTGTA	1542
Qy	438	LeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspValThrAspGlu	457
Db	1543	TTGTTTCCACCCACGATAGCTGAAGAACATTCGCTATGTCGTGAAGATGTCAACATGAT	1602
Qy	458	GluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsn	477
Db	1603	GAGATTGGAAGAAGCTGCAAGGAAGCCATGCTATGACTTTATCATGAACCTGCCTCAG	1662
Qy	478	LysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArg	497
Db	1663	AAATTTGCACCCCTGGTTGGAGAGAGAGGGGCCACGCTGAGTGGTGGGCAGAAAGCAGAGG	1722
Qy	498	IleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThr	517
Db	1723	ATCGCCATTGACGTGCCCTGGTTGCCAACCCACAGATCCTCTGCTGGACGAGGCCACG	1782
Qy	518	SerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLys	537
Db	1783	TCAGCCTTGGACACAGAAAGTGAAGCAGTGGTTTCAGTGGCTCTGGATAAGGCCAGAAAA	1842
Qy	538	GlyArgThrThrIleValAlaHisArgLeuSerThrIleArgSerAlaAspLeulle	557
Db	1843	GGTCGGACCACTTGTGATAGCTCATCGTTGTCTACGGTTTCGTAATGCCACGCTCATC	1902
Qy	558	ValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlalys	577
Db	1903	GCTGGTTTCGATGAGGATCATTTGTGGAGAAGGAAATCATGATGAGCTCATGAAGAC	1962
Qy	578	ArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIle-----	590
Db	1963	AAAGCATTTACTTCAAACTTGTCACANTGCACACGACGAGGAATGAATGAATTAGAA	2022
Qy	591	LysLysAlaAspGlu-----GlnMetGluSerMetThrTyrSerThrGluArgLys	607
Db	2023	AATCAGCTGATGAATCCAAAAAGTGAAATTGATACCTCTGGAATGCTCTTCACATGATTCA	2082
Qy	608	ThrAsnSerLeu-----ProLeuHisSerValLysSerIleLysSerAsp	622
Db	2083	GGATCCAGTCTTAATAAGAAAAAGATCCACTCGTAGGAGTGTCCGTGGATCAACAGGCCAA	2142
Qy	623	PheIleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSer	642
Db	2143	-----GACGAAAGCTTAGTACCAAGAGGCTCTGGATGAAGATATACCTCCAGTTTCC	2196
Qy	643	LeuLeuLysIleLeuLysLeuAsnLysProGluTyrProPheValValLeuGlyThrLeu	662
Db	2197	TTTTGGAGGATTATGAAGCTAAATTTAACTGAGTGGCCTTATTTTGTGTGTGTATTT	2256
Qy	663	AlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIle	682
Db	2257	TGTGCCATTAATAATGAGGTCTGCACACAGATTTTCAGATTAATATTTTCAAGATTATA	2316


```
QY 1187 LeuileValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeu 1206
Db 3724 TTAATAGTGGTGTTCAGAAATGCGAGAGTCAAGGAGCATGGCAGCATCAGCAGCTGCTG 3783

QY 1207 ArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
Db 3784 GCACAGAAAGGCATATATTTTCAATGGTCAGTGTCCAGGCT 3825

RESULT 5
ABA94365
ID ABA94365 standard; DNA; 3860 BP.
XX AC ABA94365;
XX DT 26-MAR-2002 (first entry)
XX DE Human BCRP DNA related seq Id No. 1.
XX KW Stem cell; ATP transport protein; ATP-binding cassette; antiparkinsonian;
KW hepatotropic; neurodegenerative; cytostatic; antianemic; muscular; BCRP;
KW cardiant; gene therapy; ds.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT 1..3843
XX CD CDS /*tag= a

WO200192877-A2.
PN 06-DEC-2001.
XX 30-MAY-2001; 2001WO-US17459.
XX 31-MAY-2000; 2000US-0584586.
XX 29-MAY-2001; 2001US-0866866.
XX (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX PI Sorrentino B, Schuetz J;
XX WPI; 2002-114368/15.
XX DR P-PSDB; ABB07266.
XX PT Identifying a stem cell, for treating e.g., muscular dystrophy,
XX myocardial infarction, Parkinson's disease, or neurodegenerative
XX disorders, comprises detecting the expression of an ATP transport
XX protein (BCRP) by a cell
XX PS Disclosure; Page 53-55; 87pp; English.
XX CC The invention provides a method of identifying and/or isolating a stem
XX cell that involves detecting the expression of an ATP transport protein
XX containing a conserved ATP-binding cassette (BCRP) by a cell in a sample
XX comprising stem cells. The isolated stem cells may be used in the
XX treatment of diseases such as muscular dystrophy, degenerative liver
XX disorder, myocardial infarction, Parkinson's disease, degenerative
XX disorders of the brain, and for tissue regeneration or replacement.
XX CC Haematopoietic cells can be used in bone marrow transplants (e.g., for
XX treatment of leukemia) and for ex vivo gene therapy for treating blood
XX diseases such as sickle cell anemia and thalassemia. The stem cells can
XX also be used as cell targets in gene therapy protocols. The present
XX sequence represents a sequence related to the BCRP for which no relevant
XX information has been provided in the specification.
XX SQ Sequence 3860 BP; 1135 A; 746 C; 957 G; 1022 T; 0 other;

Alignment Scores:
Pred. No.: 0 Length: 3860
Score: 3530.50 Matches: 682
Percent Similarity: 74.00% Conservative: 246
Best Local Similarity: 54.39% Mismatches: 263
Query Match: 57.32% Indels: 63
```

```
DB: 24 Gaps: 10
US-09-873-409-5 (1-1222) x ABA94365 (1-3860)

QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 151 ATGGTGGTGGGAACCTTTGGCTGCCATCATCCATGGGGCTGGACTTCTCTCTCATGATGCTG 210

QY 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
Db 211 GTGTTTGGAGAAATGACAGATATCTTTGCAAAATGCAGGAAATTTAGAAAGATCTGATGTCA 270

QY 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
Db 271 AACATCAGCTAATAGAAAGTGAATCATCATCAGGCTTCTTTCATGAATCTGGAGGAAGAC 330

QY 44 -----ArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db 331 ATGACAGATATGCTTATTATTACAGTGGAAATGGTGGGGTGTGTTGCTGCTTAC 390

QY 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db 391 ATTCAGGTTTCATTTTGGTGGCTGGCAGCTGGAAGACAAATACACAAATTTAGAAACAG 450

QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101
Db 451 TTTTTCATGCTATAATAGCAGCAGAGATAGGCTGGTTTGTATGTGCACGATGTTGGGGAG 510

QY 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db 511 CTTAACACCCGACTTACAGATGATGTCCTTAAGATTAAATGAAGGTATTGTTGACAAAATT 570

QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db 571 GGAATGTTCTTTTCAGTCAATGCGCAACATTTTTCACCTGGGTTTATAGTAGGATTACACGT 630

QY 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db 631 GGTGGAGCTAAACCTTGTGATTTTGGCCATCAGTCTCTTCTTGGACTGTCAGCTGCT 690

QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db 691 GTCTGGGCAAGATACATCTATCTTCAATTTATGATAAAGAACTCTTACCGTATGCAAAAGCT 750

QY 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db 751 GGAGCAGTAGCTGAGAGAGCTTTGGCAGCAATTAGAACTGTGATTCATTGGAGACAA 810

QY 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
Db 811 AAGAAAGAACTTGAA----- 825

QY 221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
Db 826 -----AGGTACACAAACAAATTTAGAAAGAA 849

QY 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
Db 850 GCTAAAGAAATTTGGGATAAAGAAAGCTATTACAGCCAATATTTCTATAGTGTGCTGCTTC 909

QY 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
Db 910 CTGCTGATCTATGTCATCTTATGCTCTGGCTTCTGCTGATGGGACCACTTGGTCTCTCTCA 969

QY 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
Db 970 GGGGAA-----TATTCTATTGGACAAGTACTCAGTGTATCTTTCTGTATTAAATGGG 1023

QY 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
Db 1024 GCTTTTAGTGTGGAGCAGCATCTCCAGCATTTGAGAGCATTTGCAATGCAAGAGGACCA 1083

QY 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
```


Db 1084 GCTTATGAAATCTTCAAGATAAATGATAAATAAGCCAGTAGTATTGACAGCTATTTCGAAGAGT 1143
QY 341 GlyTyrIysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
Db 1144 GGGCACAACACAGATAATATTAAAGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 1203
QY 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
Db 1204 CCATCTGAAAAGAGTTAGATCTTGAAGGGCTGAACTGGAAGTGCAGAGTGGGCAG 1263
QY 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
Db 1264 ACGTGCCCTGTTGGAACACAGTGGCTGTGGGAAGACCAACAGTCCAGCTGATGCAG 1323
QY 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
Db 1324 AGGCTCTATGACCCACAGAGGGGATCGTCAAGTGTGATGCAGCAGGATATAGGACCAT 1383
QY 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
Db 1384 AATGTAAGCTTCTACGGGAATCATTTGGTGTGGTGAAGTCAAGGATGATGATGTC 1443
QY 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspValThrAspGluGluMetGlu 460
Db 1444 ACCAGATAGCTGAAAACATTCGTATGGCGGTGAAAATGTCACCATGGATGAGATTGAG 1503
QY 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
Db 1504 AAGCTCTCAAGGAAGCAATGCTATGATCTATCATGAAGTGCCTCATAAATTGAC 1563
QY 481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
Db 1564 ACCCTGTTGGAGAGAGGGGCCAGTTGAGTGTGGGCAAGCAGAGGATCGCCATT 1623
QY 501 AlaArgAlaLeuValArgAsnProLysIleIleLeuLeuAspGluAlaThrSerAlaLeu 520
Db 1624 GCAGTGCCTCGTTGCAACCCCAAGATCTCTGCTGGATGAGGCCAGCTCAGCCCTTG 1683
QY 521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuLysAlaSerLysGlyArgThr 540
Db 1684 GACACAAAAGCAGAGCAGTGGTTGAGTGGCTCTGATGAAGCCAGAAAGGTGCGACC 1743
QY 541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560
Db 1744 ACCATTGTGATAGCTCATGTTGTTGTCTACAGTTCGTAACTGCTGACGTCATCGTGGTTTC 1803
QY 561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580
Db 1804 GATGATGGAGTCATTGTGGAGAAAGGAAATCATGATGAACCTCATGAAAGAGAAAGGCATT 1863
QY 581 TyrTyrSerLeuValMetSerGln-----AspIleLysLysAlaAsp 594
Db 1864 TACTTCAAACTGTGCACATGCACAGCAGCAGGAAATGAAGTTGAATAGAAAATGCGAGCT 1923
QY 595 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis 614
Db 1924 GATGAATCCAAAAGTGAATGTGATGCTCTGGAATGTCTTCAATGATTCAGATCCAGT 1983
QY 615 SerVal-----LysSerIleLys---SerAspPheIleAspLysAla 627
Db 1984 CTAATAAGAAAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAG 2043
QY 628 GluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLysIleLeu 647
Db 2044 CTTAGTACCAAGAGGCTCTGGATGAAGATATACCTCCAGTTTCCCTTTTGGAGGATTATG 2103
QY 648 LysLeuAsnLysProGluTyrProPheValValLeuGlyThrLeuAlaSerValLeuAsn 667
Db 2104 AAGCTAAATTAACCTGAATGGCCCTATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2163
QY 668 GlyThrValHisProValPheSerIlePheAlaLysIleIleThrMetPheGlyAsn 687

Db 2164 GGAGGGCTCCAAACCAGCATTTGCAATAATATTTCAAAAGATTATAGGGGTTTTTACAAGA 2223
QY 688 ---AsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIle 706
Db 2224 ATTGATGATCCTGAAACAAACGACAGATAAGTAAGTCTGTTTCTACTATTGTTCTAGCC 2283
QY 707 LeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGly 726
Db 2284 CTTGGAATTAATTTCTTTTATTACATTTTCTCAAGGTTTCACATTTGGCAAGCTGA 2343
QY 727 GluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIle 746
Db 2344 GAGATCCTCACAAGCGCTCCGATACATGTTTCCGATCCATGCTCAGACAGGATG 2403
QY 747 AlaTyrPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAsp 766
Db 2404 AGTTGGTTTGTGATGACCTTAAACACCATCGAGCATTCAGTACCAGGCTCGCAATGAT 2463
QY 767 IleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThr 786
Db 2464 GCTGCTCAAGTTAAAGGGCTATAGGTTCCAGGCTTGCTGTAATTACCCAGAAATAGCA 2523
QY 787 AsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyrGluMetThrPheLeuIle 806
Db 2524 AATCTTGGGACAGGAATAATATATCTTCTATCTATGTTGGCAATACACATGTTTACTC 2583
QY 807 LeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGly 826
Db 2584 TTAGCAATTTGATCCCATCATTCGAATAGCAGGAGTTGTTGAAATGAATGTTGCTGGA 2643
QY 827 PheAlaAsnLysAspLysGlnGluLysHisAlaGlyLysIleAlaThrGluAlaLeu 846
Db 2644 CRAAGCACTGAAAGATAAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATA 2703
QY 847 GluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGlu 866
Db 2704 GAAATCTTCGAACCTGTTTCTTCTGACTCAGGAGCAGAAAGTTTGAACATATGTTATGCT 2763
QY 867 GluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCys 886
Db 2764 CAGAGTTGCGAGTGCATACAGAACTCTTTGAGAAAGCACACATCTTTGGAATTACA 2823
QY 887 TyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAla 906
Db 2824 TTTTCTTCAACCCAGCAATGATGATTTTCTCTATGCTGGATGTTTCCGTTTGGAGCC 2883
QY 907 TyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIle 926
Db 2884 TACTTGGTGGCACAATAAATCACTCATGAGCTTTGAGGATGTTCTGTAGTATTTTTCAGCTGT 2943
QY 927 AlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAla 946
Db 2944 GTCTTTGGTGCATGCGCTGGGGGCAAGTCAGTTCATTGCTCTGACTATGCAAGGCC 3003
QY 947 LysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArg 966
Db 3004 AAAATATCAGCAGCCACATCATCATGATCATTTGAAAAAACCCCTTTGATTGACAGCTAC 3063
QY 967 SerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSer 986
Db 3064 AGCAGCGAGGCCCTAATGCGGAACACATTTGGAAGGAAATGTCCATTTGGTGAAGTTGTA 3123
QY 987 PhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGlu 1006
Db 3124 TTCAACTATCCACCGACCGGACATCCAGTCTCTCAGGGAAGTGGAGGTTGAGGTTGAG 3183
QY 1007 ArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGln 1026
Db 3184 AAGGGCCAGAGCTGCTGCTGGTGGGCGAGTCAGTGGCTGTGGGAAGACACAGTGGTCCAG 3243
QY 1027 LeuLeuGlnArgLeuTyrAspProValGlnGlyClnValLeuPheAspGlyValAspAla 1046
Db 3244 CTCCTGGAGCGGTTCTACGACCCCTTTGGCAGGAAAGTGTCTTGTGATGCAAGAAATA 3303

Qy	1047	LysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProVal	1066
Db	3304	AAGCGACTGAATGTTCAGTGGCTCCGAGCACACCTGGGCGCATCGTGTCCAGGAGCCCATC	3363
Qy	1067	LeuPheAsnCysserIleAlaGluAsnIleAlaTyGlyAspAsnSerArgValValPro	1086
Db	3364	CTGTTTGACTGACGACATTGCTGAGACATTGCTATGAGACACACGCCGGGTGTGTCA	3423
Qy	1087	LeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeu	1106
Db	3424	CAGGAAGAGATCGTAGGGGCGAGCAAGGAGGCGCAACATACATGCTTCATCGAGTCACTG	3483
Qy	1107	ProGluLysTyrsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLys	1126
Db	3484	CCTAATAAATATAGCACTAAAGTAGGAGACAAAGAACTCAGCTCTCTGTGGTGGCCAGAA	3543
Qy	1127	GlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGlu	1146
Db	3544	CAACGCATTGCCATTAGCTGTGGCTTGTGACAGCCCTCATATTTGCTTTTGGATGAA	3603
Qy	1147	AlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAla	1166
Db	3604	GCCACGTCAGCTCTGCATACAGAAAGTGAAGAGTGTCCAGAGGCCCTGACAAAGCC	3663
Qy	1167	ArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAsp	1186
Db	3664	AGAGAAGCCGCACTGCATTGTGATTGCTACCCTGTCCACCATCCAGATGCGAGAC	3723
Qy	1187	LeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeu	1206
Db	3724	TTAATAGTGTGTTTCAGAAATGCGAGATCAAGGAGCATGCGACGATCAGCAGCTGCTG	3783
Qy	1207	ArgAsnArgAspIleTyPheLysLeuValAsnAlaGlnSer	1220
Db	3784	GCACAGAAGGCATCTATTTTCAATGGTCAGTGTCCAGGCT	3825
RESULT 6			
AAD03504			
ID	AAD03504	standard; cDNA; 4279 BP.	
XX	AAD03504;		
XX			
DT	13-JUN-2001	(first entry)	
XX		Dog P-glycoprotein (PGP) allelic variant (Genotype A) cDNA.	
DE		Dog; P-glycoprotein allelic variant; multidrug transporter; MDR1;	
KW		drug bioavailability; transgenic animal; genetic model; ss.	
XX		Canis familiaris.	
XX			
FH	Key	Location/Qualifiers	
FT	CDS	17..3862	
FT		/*tag= a	
FT		/product= "Dog P-glycoprotein (PGP) allelic variant	
FT	allele	(Genotype A) protein"	
FT		replace (607, C)	
FT		/*tag= b	
XX			
XX	WO200123540-A2.		
XX			
PD	05-APR-2001.		
XX			
PF	28-SEP-2000; 2000WO-US26767.		
XX			
PR	28-SEP-1999; 99US-0156510.		
XX			
PA	(GENT-) GENTEST CORP.		
XX			
PI	Stocker PJ, Steinmel-crespi DT, Crespi CL, Reif TC, Patten CJ;		
XX			
DR	WPI; 2001-235373/24.		

P-PSDB; AAEO0308.

New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful for determining the bioavailability of drugs and for screening for dog PGP inhibitors -

Claim 9; Page 85-90; 11pp; English.

The invention relates to dog P-glycoprotein (PGP) also referred as multidrug transporter (MDR1) and nucleic acids encoding them. The invention also includes fragments and biologically functional variants of dog P-glycoprotein. PGP and their nucleic acids are useful for determining the bioavailability of drugs and for screening PGP inhibitors. They are useful for the diagnosis and treatment of conditions characterised by PGP activity, by reducing or increasing PGP activity in a cell. PGP nucleic acids are used as oligonucleotide probes. Complements of PGP nucleic acids are useful as antisense oligonucleotides, to induce a PGP 'knockout' phenotype. They are used to prepare a non-human transgenic animal, which are valuable as genetic models for human diseases.

The present sequence is dog P-glycoprotein (PGP) allelic variant (Genotype A) cDNA. The PGP enzyme functions as an efflux pump exporting small molecules across the cell membrane. This enzyme is a member of the ABC transporter family.

Sequence 4279 BP; 1295 A; 833 C; 1008 G; 1143 T; 0 other;

Alignment Scores:

Pred. No.:	0	Length:	4279
Score:	3530.50	Matches:	638
Percent Similarity:	73.11%	Conservative:	232
Best Local Similarity:	54.21%	Mismatches:	269
Query Match:	57.32%	Indels:	69
DB:	22	Gaps:	9

US-09-873-409-5 (1-1222) x AD03504 (1-4279)

Qy 1 MetileLeuGlyIleLeuAlaSerLeuValAsnGlyValaCysLeuProLeuMetProLeu 20
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 167 ATGTGGTGGGCAATGGCTGCCATCATCCATCGAGCTGCATCCCTCTCATGATGCTG 226
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Qy 21 ValLeuGlyGluMetSerAspAspLeuIleSerGlyCysLeu----- 34
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 227 GTTTTGGAACATGACAGATAGCTTTGCCAATTCAGAGAAATTTCAGAACAAAACTTTT 286
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Qy 35 -----ValGlnThrAsnThrTyr-SerPheArg----- 44
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 287 CCAGTTTAATAATGAAGATATTCAGAACATACACACATTTTCATCAACCATCTGGAG 346
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Qy 45 -----LeuThrLeuTyrTyrrValGlyIleGlyValAlaAlaLeuIlePhe 59
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 347 GAGGAANTGACCAGTATGCCTATTATTACATGGGATCGTGCTGGCTGGCTGGCTG 406
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Qy 60 GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg 79
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 407 GCTTACATCCAGTTTCATTCTGCTGCTGGCAGCAGGAAGACAGATACTCAAATTAGA 466
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Qy 80 LysGlnPhePheHisserValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 467 AAAACAATTTTTTCATGCTCATCGCACAGGAGATTGGCTGGTTTGACGTGATGACGTT 526
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Qy 100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 527 GGGGAGCTTAACHACCCGGCTCAGACAGATGCTCCAANAATCATGNAGNAATTGGCGAC 586
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Qy 119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 587 AAAATTGAATGTTCTTTCAATCAATAGCAACATTTTCCACCGTTTTATAGTGGGGTTT 646
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Qy 139 ValLysGlyTrpLysLeuThrLeuValThrLeuSerThr-SerProLeuIleMetaLaser 158
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 647 ACAGTGGTTTGAAGCTAAACCCCTTTGATTTTGGCCATCAGCCCTGTTCTTGACATTCA 706
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||


```
Db 2774 TACATGATGCACAGAGTTTGCAGGTACCATACAGAAACTCTTTGAGGAAGACACATC 2833
QY 883 ILeGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPhe 902
Db 2834 TTCGGGGTCTCATTTTCTATCACCAGGCATATGATATTTTCTATGCTGCTGTTTC 2893
QY 903 ArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleVal 922
Db 2894 CGGTTTGGTCTACTTGGTGGCAATAGAGTTTCATGAACCTTTCAGGATGTTCTTTGGTA 2953
QY 923 PheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGlu 942
Db 2954 TTTCTAGCTATTCTTTGGTGCATGGCAGTGGCAGGTCAGTTTCATTGCTCCTGAC 3013
QY 943 TyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGlyLysProAsn 962
Db 3014 TATGCCAAGCCAAAGTATCAGAGCCAGCTCATATGATCATTTGAAAAGCCCTCTG 3073
QY 963 IleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPhe 982
Db 3074 ATTGACAGCTACAGACCTCAGCGCCTCAAGCCAAATACGTTGGAAGAAATGTGACATTT 3133
QY 983 ArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSer 1002
Db 3134 AATGAGGTGCTGTTCACATATCCACATCGACACAGATCCCGCTGCTCCAGGGGCTGAGC 3193
QY 1003 LeuSerIleGluArgGlyLysThrValAlaPheValIleGlySerSerGlyCysGlyLysSer 1022
Db 3194 CTCGAGGTGAGAGGGCCAGAGCTGGCCCTCTGTAGTAGAGTGGCTGTGGAGAGAGC 3253
QY 1023 ThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlnValLeuPheAsp 1042
Db 3254 ACAGTTGTTTCAGCTCTAGAGCGCTTCTATGACCCCTTGCTGCTGCTTAATTGAT 3313
QY 1043 GlyValAspAlaLysGlnLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValPro 1062
Db 3314 GGCAAGAGATAAGACCTGAATGTCTCAGTGGCTCGAGCACACCTGGGCATCGTGTCT 3373
QY 1063 GlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSer 1082
Db 3374 CAGAGCCCATCTCTGTTGACTCAGATATGCCGAGAAATTCCTTATGAGACAACAGC 3433
QY 1083 ArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPhe 1102
Db 3434 CGGGTCTATCATGAGAGATTATCGAGCGAGCCAGGAGGCCAATACATACACCTTC 3493
QY 1103 IleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSer 1122
Db 3494 ATCGAGACACTCCCTCAGAAATACACACAGAGTAGGAGACAAAGAACCCAGCTCTCT 3553
QY 1123 GlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuGlnLysProLysIleLeu 1142
Db 3554 GGTGGCCAGAAACAGCGCATGTGCATAGCTCGCGCTCTTTAGACAGCTCATATTTTG 3613
QY 1143 LeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAla 1162
Db 3614 CTTTGGATGAGCTATCATCAGCTCTGGATACAGAAAGTGCAAGGTGTGTCGAAGGCC 3673
QY 1163 LeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIle 1182
Db 3674 CTGGCAAAAGCCAGAGAAGGCGCACCTGTCATTGTGATCGCCACCGCTTGTCCACCATC 3733
QY 1183 GlnAsnAlaAspIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHis 1202
Db 3734 CAGAATGCAGATTAAATAGTGGTGTTCAGAAATGGCAAGTCAAGGAGCATGGCACAT 3793
QY 1203 GlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
Db 3794 CAACAGCTGCTGCCCAAGAAAGCACTATTTTCCATGTGTAGTTCAGGCT 3847
```

RESULT 7
AAD03505

```
ID AAD03505 standard; cDNA; 4279 BP.
AC AAD03505;
DT 13-JUN-2001 (first entry)
DE Dog P-glycoprotein (PGP) allelic variant (Genotype B) cDNA.
KW Dog; P-glycoprotein allelic variant; PGP; multidrug transporter;
KW MDRI; drug bioavailability; transgenic animal; genetic model; ss.
OS Canis familiaris.
FH Key Location/Qualifiers
FT CDS 17..3862
FT /*tag= a
FT /product= "Dog P-glycoprotein (PGP) allelic variant
FT (Genotype B) protein"
FT replace (91, f)
FT /*tag= b
FT allele replace (607, C)
FT /*tag= c
XX WO200123540-A2.
PN 05-APR-2001.
XX 28-SEP-2000; 2000WO-US26767.
XX 28-SEP-1999; 99US-0156510.
XX (GENT-) GENTEST CORP.
XX Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;
XX WPI; 2001-235373/24.
XX P-PSDB; AAE00309.
XX New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
XX for determining the bioavailability of drugs and for screening for dog
XX PGP inhibitors -
PS Claim 9; Page 93-99; 111pp; English.
XX The invention relates to dog P-glycoprotein (PGP) also referred
XX as multidrug transporter (MDRI) and nucleic acids encoding them.
XX The invention also includes fragments and biologically functional
XX variants of dog P-glycoprotein. PGP and their nucleic acids are
XX useful for determining the bioavailability of drugs and for
XX screening PGP inhibitors. They are useful for the diagnosis and
XX treatment of conditions characterised by PGP activity, by
XX reducing or increasing PGP activity in a cell. PGP nucleic acids
XX are used as oligonucleotide probes. Complements of PGP nucleic
XX acids are useful as antisense oligonucleotides, to induce a PGP
XX 'knockout' phenotype. They are used to prepare a non-human
XX transgenic animal, which are valuable as genetic models for
XX human diseases.
XX The present sequence is dog P-glycoprotein (PGP) allelic variant
XX (Genotype B) cDNA. The PGP enzyme functions as an efflux pump
XX exporting small molecules across the cell membrane. This enzyme
XX is a member of the ABC transporter family.
XX Sequence 4279 BP; 1296 A; 833 C; 1008 G; 1142 T; 0 other;
SQ
Alignment Scores:
Pred. No.: 0 Length: 4279
Score: 3530.50 Matches: 682
Percent Similarity: 73.13% Conservative: 238
Best Local Similarity: 54.21% Mismatches: 269
Query Match: 57.32% Indels: 69
DB: 22 Gaps: 9
US-09-873-409-5 (1-1222) x AAD03505 (1-4279)
```



```

Db 2234 ATCTTTACCGAGATGAGATCTCTGAAACAAACGACAGATAGTAAACATGTTTCTGTGA 2293
      :::::  :::::  |||  :::::  :::::  :::::  :::::  :::::  :::::
Qy 703 IlePheValIleLeuGlyValIleCysPheValSerPheMetGlnGlyLeuPheTyr 722
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 2294 TTGTTTCTAGTCTTGGAATATTCTTTTATTACATTTTCTCCAGGCGCTTCACATTT 2353
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Qy 723 GlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheIlysAlaMetLeu 742
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 2354 GGCAAAGCTGGGAGATCTCTACATAAGCGCGCTTCGATACATGGTTTTTCAGATCCATGCTG 2413
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Qy 743 TyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrIle 762
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 2414 AGACAGATGTCAGCTGGTTTGATGACCTTAAACACCACTGGACATTCACACAGG 2473
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Qy 763 LeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThr 782
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 2474 CTTGCCAATGATCGCGCTCAAGTTAAAGGGCTATAGTTCCAGGCTGCTGTCATTACC 2533
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Qy 783 GlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMet 802
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 2534 CAGAATATAGCAAACTCTGGGACAGGCATTTATATATCTTAACTATGCTTGGCAATTA 2593
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Qy 803 ThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrClyMetIleGluThrAla 822
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 2594 ACATTTTACTCTTGCAATTTGACCATTTGCAATAGCAGGAGTTGTTGAATGAAA 2653
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Qy 823 AlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAla 842
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 2654 ATGTTGCTGGCAACGACTGAAGATAGAAAGAGCTAGAGGAGCTGGGAAGATTGCT 2713
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Qy 843 ThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGlu 862
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 2714 ACAGAAGCCATGAAAACTTCGAACTGTGTCTTTGACTCGGAGCAGAAAGTTTGA 2773
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Qy 863 GlnMetTyrGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIle 882
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 2774 TACATGTATGCAAGAGTTTGCAGATGCCATACAGAAACTCTTTGAGGAAAGCACACATC 2833
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Qy 883 IleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPhe 902
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 2834 TTCCGGGTCTCATTTCTATCACCCAGGCATATGATGATTTTCTATGCTGCTGTTTC 2893
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Qy 903 ArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleVal 922
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 2894 CGGTTTGGTGCCTACTTGGTGGCAATGAGTTCTATGAACCTTTCAGGATGTTCTTTGGTA 2953
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Qy 923 PheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGlu 942
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 2954 TTCTCAGCTATTGCTTTGTTGTCATGGCAGTGGGCGAGTTCAGTTTCATTGCTCCTGAC 3013
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Qy 943 TyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuLysLysProAsn 962
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 3014 TATGCCAAAGCCAAAGTATCAGACGCCACGTCATCATGATCAITGAAAAGCCCTCTG 3073
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Qy 963 IleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPhe 982
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 3074 ATTGACAGCTACAGCCCTCACGSCCTCAAGCCAAATACGTTGGAAGAAATGTGACATTT 3133
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Qy 983 ArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSer 1002
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 3134 AATGAGGTGCTGTTCAACTATCCACTCGACACAGACATCCCCGCTGCCAGGGCGCTGAGC 3193
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Qy 1003 LeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSer 1022
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 3194 CTCGAGTGAAGAAGGGCAGACGCTGGCCCTCGTAGGTAGCAGTGGCTGTGGGAAGAGC 3253
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Qy 1023 ThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAsp 1042
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 3254 ACAGTTGTTTCAGCTCTAGAGCGCTTCTATGACCCCTTGCTGCTGCTCAGTGTAAATGAT 3313
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Qy 1043 GlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValPro 1062
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::

```

```

Db 3314 GGCAAAGAGATAAGCACTGAATGTCCAGTGGCTCCGAGCACACCTGGGCATCGTGCT 3373
Qy 1063 GlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSer 1082
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 3374 CAGGAGCCCATCTGTTTGTACCTGCAGCATTCGCGAACAATTTGCCTTATGGAGACAACAGC 3433
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Qy 1083 ArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPhe 1102
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 3434 CGGGTCGTATCATCATGAAGAGATTATGCGAGCGCCCAAGGAGGCCCAACATACACCACTTC 3493
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Qy 1103 IleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSer 1122
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 3494 ATCGAGACACTCCCTCGAGAAATACACACCAGAGTAGGAGACAAAGAACCCAGCTCTCT 3553
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Qy 1123 GlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeu 1142
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 3554 CGTGGCCAGAAACAGCGCATTTGCCATAGCTCGCGCTCTTTGTAGACAGCCTCATATTTTG 3613
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Qy 1143 LeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAla 1162
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 3614 CTTTTGGATGAAGCTACATCAGCTCTGGATACAGAAAGTGAAGGTTGTCCAAGAAGCC 3673
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Qy 1163 LeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIle 1182
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 3674 CTGGACAAAGCCAGAGAGCGCCGACCTCGATTGTGATCGCCACCGCTTGCCACCATC 3733
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Qy 1183 GlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHis 1202
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 3734 CAGAATGCAGATTTAATAGTGTGTTTCAGATGGCAAGTGCAGAGCATGCGACACAT 3793
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Qy 1203 GlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 3794 CAACAGCTGCTGCCAGAAAGGCATCTATTTTCCATGTCAGTGTCCAGGCT 3847
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::

```

RESULT 8

AD03488
ID AD03488 standard; cDNA; 4279 BP.

XX AD03488;

XX 13-JUN-2001 (first entry)

XX Dog P-glycoprotein (PGP) cDNA #1.

XX Dog; P-glycoprotein; PGP; multidrug transporter; MDR1;
KW drug bioavailability; transgenic animal; genetic model; ss.

XX Canis familiaris.

XX Key Location/Qualifiers
FH 17..3862
CDS /*tag= a

FT /product= "Dog P-glycoprotein (PGP) #1"

XX WO200123540-A2.

XX 05-APR-2001.

XX 28-SEP-2000; 2000WO-US26767.

XX 28-SEP-1999; 99US-0156510.

XX (GENT-) GENTEST CORP.

XX Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;

XX WPI; 2001-235373/24.

XX P-PSDB; AA00303.

XX New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
PT for determining the bioavailability of drugs and for screening for dog
PT PGP inhibitors -

XX

Claim 3; Page 58-63; 111pp; English.

PS The invention relates to dog P-glycoprotein (PGP) also referred
 XX as multidrug transporter (MDR1) and nucleic acids encoding them.
 CC The invention also includes fragments and biologically functional
 CC variants of dog P-glycoprotein. PGP and their nucleic acids are
 CC useful for determining the bioavailability of drugs and for
 CC screening PGP inhibitors. They are useful for the diagnosis and
 CC treatment of conditions characterized by PGP activity, by
 CC reducing or increasing PGP activity in a cell. PGP nucleic acids
 CC are used as oligonucleotide probes. Complements of PGP nucleic
 CC acids are useful as antisense oligonucleotides, to induce a PGP
 CC 'knockout' phenotype. They are used to prepare a non-human
 CC transgenic animal, which are valuable as genetic models for
 CC human diseases.
 CC The present sequence is dog P-glycoprotein (PGP) cDNA. This
 CC sequence is also referred as genotype C cDNA. The
 CC PGP enzyme functions as an efflux pump exporting small molecules
 CC across the cell membrane. This enzyme is a member of the ABC
 CC transporter family.

XX Sequence 4279 BP; 1294 A; 834 C; 1008 G; 1143 T; 0 other;
 SQ

Alignment Scores:

Pred. No.: 0 Length: 4279
 Score: 3525.50 Matches: 681
 Percent Similarity: 73.05% Conservative: 238
 Best Local Similarity: 54.13% Mismatches: 270
 Query Match: 57.24% Indels: 69
 DB: 22 Gaps: 9

US-09-873-409-5 (1-1222) x AAD03488 (1-4279)

Qy 1 MetLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
 Db 167 ATGTTGGTGGGACAATGGCTGCATCATCCATGGAGCTGCACCTCTCATGATCGT 226
 Qy 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu----- 34
 Db 227 GTTTTGGAAACATGACAGATAGCTTTGCAATGCAGGAATTTCAAGAACAAACTTTT 286
 Qy 35 -----ValGlnThrAsnThrTyrSerPhePheArg----- 44
 Db 287 CCAGTTATAATTAAGAGATTATACGAACAATACACAACTTTTCATCAACCATCTGGAG 346
 Qy 45 -----LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe 59
 Db 347 GAGGAATGACCCAGTATGCTATTATTACAGTGGGATCGGTGCTGGTGGCT 406
 Qy 60 GlyTyrIleGlnIleSerLeuTyrIleThrAlaAlaArgGlnThrLysArgIleArg 79
 Db 407 GCTTACATCCAGTTTCACTTCTGGTGGCTGGCAGGAGACAGATACCAAAATTAGA 466
 Qy 80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIle 99
 Db 467 AAACAATTTTTCATGCTATCATCGACAGGAGATGGCTGGTTGACGTGCATGACGTT 526
 Qy 100 GlyGluLeuAsnThrArgMetThr-----AspIleAspIleSerAspGlyIleGlyAsp 118
 Db 527 GGGGAGCTTAAACCCGCTCACAAGACGATGCTCCAAAATCAATGAAGGAATTTGGCGAC 586
 Qy 119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
 Db 587 AAAATTGGAATGTTCTTCACTCAATAGCAACATTTTTCACCGGTTTATAGTGGGTTT 646
 Qy 139 ValLysGlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
 Db 647 ACAGTGGTTGGAAGCTAACCTTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTTCA 706
 Qy 159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
 Db 707 GCCGCCCATCTGGGCAAGATACTATCTTCTTACTATGAAGAACTCTTGGCCCTATGCA 766

Qy 179 LysAlaGlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArg 198
 Db 767 AAAGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGAACTGTGATTGCTTTGGA 826
 Qy 199 AlaGlnGlyLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTyrPhe 218
 Db 827 GGCACAAAGAAAGAACTTGA----- 847
 Qy 219 TyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeu 238
 Db 848 -----AGGTACAACAAAATTTA 865
 Qy 239 LysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAla 258
 Db 866 GAAGAAAGCTAAAGGAATTTGGGATAAAGAAAGCTATCACGCCCAACATTTCTATTGGTGC 925
 Qy 259 ValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheThrTyrGlyThrSerLeu 278
 Db 926 GCTTTCTTATTGATCTATGATCATATGCTCTGGCTTTCTGGTATGGACCTCTTGGTC 985
 Qy 279 LeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIle 298
 Db 986 CTCCTCAGTGAA-----TATTCTATGGACAAGTACTCACTGTCTCTTTCTGTATTA 1039
 Qy 299 HisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArg 318
 Db 1040 ATTTGGGGCTTTTATGATTATGGACAGGCATCCCCAAGCATTTGAAGCATTTGCAACGCAAGA 1099
 Qy 319 GlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSer 338
 Db 1100 GGAGCAGCTTATGAATCTTCAAGATAATTGACAATAAACCACCAAGCATTTGACACTATTCG 1159
 Qy 339 ThrAlaGlyTyrLysProGluSerIleGlyThrValGluPheLysAsnValSerPhe 358
 Db 1160 AAGAGTGGACATTAACAGATATATTAAGGAAATTTGGAATTCAAAATGTTCACTTC 1219
 Qy 359 AsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSer 378
 Db 1220 AGTTACCTTCTCGAAAAGAAAGTAAAGATCTTAAAGGTCCTCAACCTGAAGGTTCAAGT 1279
 Qy 379 GlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeu 398
 Db 1280 GGGCAGACAGTGGCTGGTGGGAAACAGTGGCTGGGGAAGAGACGACCGTGCAGCTG 1339
 Qy 399 LeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArg 418
 Db 1340 ATGCAGAGGCTCTATGACCCACAGATGGATGGTCTGTATTGATGGACAGGACATTAGG 1399
 Qy 419 AlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeu 438
 Db 1400 ACCATAAATGTAAGGCATCTTCGGGAAATTTACTGTGTGTGTGAGTCAGGAGCCTGTGTTG 1459
 Qy 439 PheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGlu 458
 Db 1460 TTTGCCACACAGATAGCTGAACCAATTCGCTATGGCCGCAAAATGTCACCATGATGAG 1519
 Qy 459 MetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLys 478
 Db 1520 ATTGAGAAAGCTGTTAAGGAAGCCATTCCTATGATTTTATCATGAACATACCTAATAAA 1579
 Qy 479 PheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIle 498
 Db 1580 TTTGACACTCTGGTTGGAGAGAGAGGGGCCAGCTGAGTGGTGAGTCAGGAGCCTGTGTTG 1639
 Qy 499 AlaIleAlaAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSer 518
 Db 1640 GCCATTGCTCGGCCCTTGGTTTCGCAACCCCAAGATTTCTCTGTGGATGAGGCAACGTCA 1699
 Qy 519 AlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGly 538
 Db 1700 GCTCTGACACTGAAAGTGAAGCAGTGGTTCAGTGGCCCTGGATAGGCCAGAAAGGC 1759
 Qy 539 ArgThrThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuVal 558

|||||
Db 1760 CGGACTACCATTTGATAGTCATCGTTGCTACAGTTGCGTAATCCGATGTCATTGCT 1819
Qy 559 ThrLeuLysAspGlyMetLeuAlaGluLysGlyValAlaHisAlaGluLeuMetAlaLysArg 578
|||||
Db 1820 GGTTTTGATGATGAGTCACTTGGTGGAGAAGGAATCATCATGCAACTCATGAAGAAGAG 1879
Qy 579 GlyLeuTyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGlu 598
|||||
Db 1880 GGCATTTACTTCAAACTTGTCAAAATGCAG--ACAAGAGGAAATGAAATTCAGTTAGAA 1936
Qy 599 SerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSer 618
|||||
Db 1937 AATGCCACTGGTGAATCCAAAGTGAAGTGAATGCTTGGAAATG---TCTCAAAGAT 1993
Qy 619 IleLysSerAspPheIle----- 624
Db 1994 TCAGGGTCCAGTTTAATAAAGAAAGATCAACTCGCAGGAGTATACATGCACCACAAGGC 2053
Qy 625 ----AspLysAlaGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeu 643
|||||
Db 2054 CAAGACAGAAAGCTTGGTCAAAAGAGGACTTGAATCAGAAATGTACCTCCAGTTCCCTTC 2113
Qy 644 LeuLysIleLeuLeuLeuLysProGluTyrProPheValValLeuGlyThrLeuAla 663
|||||
Db 2114 TGGAGATTTCTGAAGCTGAACCTCAACTGAATGGCCTTATTTTGTGGTATATTTTGT 2173
Qy 664 SerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleThr 683
|||||
Db 2174 GCTATTTAATAACGGAGGCTGCACACGACATTTTCAATATATATTTCAAGGATATAGGG 2233
Qy 684 MetPhe---GlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMet 702
|||||
Db 2234 ATCTTTACCAGATCAGGATCTCTGAAAACAAAACGACAGAAATAGTAACATGTTTCTGTA 2293
Qy 703 IlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyr 722
|||||
Db 2294 TTGTTTCTAGTCTCGGAATATTTCTTTTATACATTTTCTCCAGGGCTTTCACATTT 2353
Qy 723 GlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeu 742
|||||
Db 2354 GGCAAAGCTGGGAGATCTCTCAAGCGGCTTCGATACATGTTTTCAGATCCATGCTG 2413
Qy 743 TyrGlnAspIleAlaTyrPheAspGluLysGluAsnSerThrGlyGlyLeuThrIle 762
|||||
Db 2414 AGACAGGATGTCAAGTGTGTGATGACCTTAAACCAACCACTGGAGCATTTGACAACACG 2473
Qy 763 LeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThr 782
|||||
Db 2474 CTGGCAATGATCGGCTCAAGTTAAAGGGGCTATAGTTCCAGGCTTGTGTCATTACC 2533
Qy 783 GlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyrTpGluMet 802
|||||
Db 2534 CAGATATAGCAAACTTGGACAGGCATTTATATCTTAATCTATGTTGGCAATTA 2593
Qy 803 ThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAla 822
|||||
Db 2594 ACATTTTACTCTTACCAATTTGACCATCATTTGCAATAGCAGAGCTTGTGAAATGAAA 2653
Qy 823 AlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAla 842
|||||
Db 2654 ATGTTCTCTGGCAAGCACTGAAAGATAAGAAAGAGCTAGAAAGAGCTGGAAGATTGCT 2713
Qy 843 ThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGlu 862
|||||
Db 2714 ACAGAAAGCCATCGAAACCTTCGAACTGTGTGTTCTTTGACTCGGGAGCAGAGTTTGA 2773
Qy 863 GlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIle 882
|||||
Db 2774 TACATGTATGCACAGAGTTTGCAGATACCATACAGAAATCTCTTGGAGAAAGCACACATC 2833
Qy 883 IleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPhe 902
|||||

2834 TTCGGGGTCTCTATTTTCTATATCACCAGGCAATGATGATATTTTCTATGCTGGCTGTTTC 2893
Qy 903 ArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleVal 922
|||||
Db 2894 CGGTTTGGTGCCTACTTGGTGCAAAATGATGTTCAAGAACTTTTCAGGATGTTCTTTTGGTA 2953
Qy 923 PheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGlu 942
|||||
Db 2954 TTTTCAGCTATTGTTCTTGGTGCCATGGCAGTGGGCGAGTCAGTTTCATTGCTCTCTGAC 3013
Qy 943 TyrSerLeuAlaLysSerGlyAlaAlaHisLeuPheAlaLeuGluLysLysProAsn 962
|||||
Db 3014 TATGCCAAAGCCAAAGTATCATGAGCCCACTCATCATCATCATGATCATGATAAAGACCCCTCTG 3073
Qy 963 IleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPhe 982
|||||
Db 3074 ATTGACAGCTACAGCCCTCAGCGCCTCAAGCCCAATACGTTGGAAGAAATGTGACATTT 3133
Qy 983 ArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSer 1002
|||||
Db 3134 AATGAGGTGCTGTTCAACTATCCCACTCGACACGACATCCCGTGTCTCCAGGGGCTGAGC 3193
Qy 1003 LeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSer 1022
|||||
Db 3194 CTCGAGGTGAAGAGGCCAGACGCTGGCCCTCGTAGGTAGCAGTGGCTGTGGAAAGAGC 3253
Qy 1023 ThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlnValLeuPheAsp 1042
|||||
Db 3254 ACAGTTGTTTCAGTCTCTAGAGCGCTTCTATGACCCCTTGGCTGGTTCAGTGTCTAATTGAT 3313
Qy 1043 GlyValAspAlaLysGluLeuAsnValGlnTyrLeuArgSerGlnIleAlaIleValPro 1062
|||||
Db 3314 GGAAGAGATAAAGCACCTGAATGTCCAGTGGCTCCGAGCACACCTGGGCATCGTGTCT 3373
Qy 1063 GlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSer 1082
|||||
Db 3374 CAGGAGCCCATCTGTTGACTGACGACATTCGCGAACAATTCCTTCATGGAGAACACAGC 3433
Qy 1083 ArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPhe 1102
|||||
Db 3434 CGGTCGTATCATCATGAGAGATTTATGAGGAGCCAGGAGGCCAATACACCACTTC 3493
Qy 1103 IleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSer 1122
|||||
Db 3494 ATCGAGACACTCCCTGAGAAATACAAACACAGATAGGAGACAAAGGAACCCAGCTCTCT 3553
Qy 1123 GlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuGlnLysProLysIleLeu 1142
|||||
Db 3554 GGTGGCCAGAAACAGCGCATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3613
Qy 1143 LeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAla 1162
|||||
Db 3614 CTTTGGATGAGCTACATCAGCTCTGGATACAGAAAGTGAAGAGTTGTTCCAGAGGCC 3673
Qy 1163 LeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIle 1182
|||||
Db 3674 CTGGCAAAAGCCAGAGAGCGCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3733
Qy 1183 GlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHis 1202
|||||
Db 3734 CAGAATGCGATTTAATAGTGTGTTTCAGAATGGCAAGTCAAGGAGGATGGCACACAT 3793
Qy 1203 GlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
|||||
Db 3794 CAACAGCTGCTGCCAGAAAGGCATCTATTTTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCT 3847
RESULT 9
AAD03489
ID AAD03489 standard; cDNA; 4317 BP.
XX
AC AAD03489;
XX
DT 13-JUN-2001 (first entry)


```

XX Dog P-glycoprotein (PGP) cDNA #2.
XX
XX DE
XX
XX Dog; P-glycoprotein; PGP; multidrug transporter; MDR1;
XX drug bioavailability; transgenic animal; genetic model; ss.
XX
XX Canis familiaris.
XX
XX OS
XX Location/Qualifiers
XX Key 70..3912
XX CDS /*tag= a
XX /product= "Dog P-glycoprotein (PGP) #2"
XX
XX WO200123540-A2.
XX
XX
XX 05-APR-2001.
XX
XX 28-SEP-2000; 2000WO-US26767.
XX
XX 28-SEP-1999; 99US-0156510.
XX
XX (GENT-) GENTEST CORP.
XX
XX Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;
XX
XX WPI: 2001-235373/24.
XX P-PSDB; AAE00304.
XX
XX New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
XX for determining the bioavailability of drugs and for screening for dog
XX PGP inhibitors -
XX
XX Claim 1; Page 66-72; 111pp; English.
XX
XX The invention relates to dog P-glycoprotein (PGP) also referred
XX as multidrug transporter (MDR1) and nucleic acids encoding them.
XX The invention also includes fragments and biologically functional
XX variants of dog P-glycoprotein. PGP and their nucleic acids are
XX useful for determining the bioavailability of drugs and for
XX screening PGP inhibitors. They are useful for the diagnosis and
XX treatment of conditions characterised by PGP activity, by
XX reducing or increasing PGP activity in a cell. PGP nucleic acids
XX are used as oligonucleotide probes. Complements of PGP nucleic
XX acids are useful as antisense oligonucleotides, to induce a PGP
XX 'knockout' phenotype. They are used to prepare a non-human
XX transgenic animal, which are valuable as genetic models for
XX human diseases.
XX
XX The present sequence is dog P-glycoprotein (PGP) cDNA. The
XX PGP enzyme functions as an efflux pump exporting small molecules
XX across the cell membrane. This enzyme is a member of the ABC
XX transporter family.
XX
XX SQ Sequence 4317 BP; 1293 A; 844 C; 1019 G; 1161 T; 0 other;
XX
XX Alignment Scores:
XX Pred. No.: 0 Length: 4317
XX Score: 3525 50 Matches: 680
XX Percent Similarity: 73.05% Conservative: 239
XX Best Local Similarity: 54.05% Mismatches: 270
XX Query Match: 57.24% Indels: 69
XX DB: 22 Gaps: 9
XX
XX US-09-873-409-5 (1-1222) x AAD03489 (1-4317)
XX
XX QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
XX |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX 217 ATGTGGTGGGACAAATGGCTGCATCCATCGAGCTGCATCCCTCTCATGATGCTG 276
XX
XX QY 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu----- 34
XX |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX 277 GTTTTGGAAACATGACAGATAGCTTTGGCAAAATTCAGGAATTTCAAGAAACAAACTTTT 336
XX
XX QY 35 -----ValGlnThrAsnThrTyrSerPheArg----- 44

```

```

Db 337 CCAGTTTATAATTAATGAAAGTATTACGAAACAATACACAACATTTTCACACCATCTGGAG 396
QY 45 -----LeuThrLeuTyrValGlyIleGlyValAlaAlaLeuPhe 59
Db 397 GAGGAAATGACCAGTATGCCCTATTATTACAGTGGATCGGTGCTGGTGGTGGT 456
QY 60 GlyTyrIleGlnIleSerLeuTyrIleThrAlaAlaArgGlnThrLysArgIleArg 79
Db 457 GCTTACATCCAGTTTCATTCTGGTGGCAGCAGGAGGAGACAGATACTCAAAATTAGA 516
QY 80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIle 99
Db 517 AAACAATTTTTCATGCTATCGACAGAGATGGCTGTGGTTCGACGTGATGACGTT 576
QY 100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
Db 577 GGGGAGCTTAACACCCGGCTCACAGACATGTCTCCAAATCAATGAAGAAATGGCGAC 636
QY 119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
Db 637 AAAGTTGGATGTTCTTTCAATCAATAGCAACATTTTTCACCGTTTATAGTGGGTTT 696
QY 139 ValLysGlyTyrPheLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
Db 697 ACACCTGGTTGGAAGCTAACCCCTTGTGATTTTGGCCATCAGCCCTGTTCTTGACATTCA 756
QY 159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
Db 757 GCGCCATCTGGGCAAAAGATACTATCTTCAATTAAGAAAGTCTTGGCTCTATGCA 816
QY 179 LysAlaGlyValAlaGluValLeuSerSerIleArgThrValIleAlaPheArg 198
Db 817 AAAGCTGGAGCAGTAGCTGAAGAGCTTAGCAGCAATCAGAACTGTGATGCTTTGGA 876
QY 199 AlaGlnGluGlyGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPhe 218
Db 877 GGACAAAGAAAGAACTTGAA----- 897
QY 219 TyrPheProGlnTrpLeuLeuSerCysValLeu---PheValArgTyrThrGlnAsnLeu 238
Db 898 -----AGGTACAAACAAAATTTA 915
QY 239 LysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAla 258
Db 916 GAAGAAGCTAAAGAATTGGGATAAAGAAAGCTATCACGCCCAACATTCTATTGGTGCC 975
QY 259 ValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIle 278
Db 976 GCTTTCTTATTGATCATGATCATATGCTCTGGCTTTCTGGTATGGGACCTCTTGTC 1035
QY 279 LeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPheSerValIle 298
Db 1036 CTCTCCAGTGA-----TATACTATTGGACAGGTACTCACTGCTCTTCTTTCTGATTA 1089
QY 299 HisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArg 318
Db 1090 ATTGGGGCTTTTATGATTGGACAGGCATCCCAAGCATTTGAAGCATTTGCAACGCAAGA 1149
QY 319 GlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSer 338
Db 1150 GGAGCAGCTTATGAAATCTCAAGATAATTCACATAAACCACCAAGCATTTGACACTATTCG 1209
QY 339 ThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPhe 358
Db 1210 AAGAGTGGACATAAACACAGATATATTAAGGGAAATTTGGAATTCAAAATGTTCATCTC 1269
QY 359 AsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSer 378
Db 1270 AGTTACCTTCTCGAAAGAGAGTTAGATCTTAAGGGTCTCAACCTGAAGGTTCAGAGT 1329
QY 379 GlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeu 398

```


Db 1330 GGGCAGACAGTGGCGTGGTGGGAACAGTGGCTGGGGAAGACAGCAGCCGTGCAGCTG 1389
Qy 399 LeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValaspGluAsnAspIleArg 418
Db 1390 ATGCAGAGGCTCTATGACCCACAGATGCGATGGTCTGTATTGATGGACAGCAGCATAGG 1449
Qy 419 AlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeu 438
Db 1450 ACCATAAATGTAAGGCATCTTCGGGAATATTACTGTGTGTGGTGGTGGTGGTGGTGGT 1509
Qy 439 PheGlyThrThrIleSerAsnAsnIleLeuTyrGlyArgAspAspValThrAspGluGlu 458
Db 1510 TTTCGCCACCATGATGCTGAACATCTCTATGCGCGGAAATGCTACCATGGATGAG 1569
Qy 459 MetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLys 478
Db 1570 ATTGAGAAAGCTGTTAAGGAGCCATGCTATGATTTTATCATGAACCTACCTAATAA 1629
Qy 479 PheAsnThrLeuValGlyGlnLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIle 498
Db 1630 TTTGACACTCTGGTGGAGAGAGAGGGCGCGCTGAGTGGTGGACAGAAACAGAGAATC 1689
Qy 499 AlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSer 518
Db 1690 GCCATTGCTCGGCGCTGGTTCGCAACCCAGAAATTCCTCTGCTGATGAGGCAACGTCA 1749
Qy 519 AlaLeuAspSerGlnSerLysSerAlaValGlnAlaAlaLeuGlnLysAlaSerLysGly 538
Db 1750 GCTCTGGACATGAAGTGAAGCAGTGGTTCAGTGGCCCTGGATGAAGCCAGAAAGGC 1809
Qy 539 ArgThrThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspIleVal 558
Db 1810 CGGACTACCATTTGATAGTATCATCTGTTGCTACAGTTTCGTAATCCGATGTCATTGCT 1869
Qy 559 ThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArg 578
Db 1870 GGTGTTGATGGATGATCATTTGGTGAAGAAGGAAATCATGATGAATCAAGAAAGAG 1929
Qy 579 GlyLeuTyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGlu 598
Db 1930 GGCAATTTACTTCAAACTGTCACAAATGCAG---ACAAGAGGAATGAATGATGTTAGAA 1986
Qy 599 SerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSer 618
Db 1987 AATGCCACTGGTGAATCCAAAGTGAAGATGATGCTTGGAAATG---TCTCCAAAGAT 2043
Qy 619 IleLysSerAspPheIle----- 624
Db 2044 TCAGGCTCCAGTTTATAAAGAAAGATCAACTCGCAGGAGTATACATGCACCACAGGC 2103
Qy 625 ---AspLysAlaGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeu 643
Db 2104 CAAACAGAAAGCTGGTCAAAAGAGGACTGGAATGAGATGTTCTTCAGTTTCTTTC 2163
Qy 644 LeuLysIleLeuLysLeuAsnLysProGluTyrProPheValValLeuGlyThrLeuAla 663
Db 2164 TGGAGGATTTCTGAAGCTGAACCTCACTGAATGCGCTTATTTTGTGGTATATTTGT 2223
Qy 664 SerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThr 683
Db 2224 GCTATTATAACCGAGCGCTGCAACCGCATTTTCAATATATTTTCAAGGATATAGGG 2283
Qy 684 MetPhe---GlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMet 702
Db 2284 ATCTTTACCGAGATGAGGATCTCGAAACAAACGACAGAAATAGTAACATGTTTCTGTA 2343
Qy 703 IlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyr 722
Db 2344 TTGTTTCTAGTCTCTGGAATATTCTTTTATATACATTTTCTCCAGGCGCTTCACATT 2403
Qy 723 GlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeu 742
Db 2404 GGCAAGCTGGGAGATCTCTCACTAAGCGGCTTCGATACATGTTTTCAGATCCATGCTG 2463

Qy 743 TyrGlnAspIleAlaTyrPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIle 762
Db 2464 AGACAGGATGTGAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2523
Qy 763 LeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThr 782
Db 2524 CTTGGCAATGATGCGGCTCAAGTTAAAGGGGCTATAGGTTCAGGCTTCTGTGTCATTAC 2583
Qy 783 GlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyrGluMet 802
Db 2584 CAGATATAGCAAAATCTTGGGACAGGCATTATATATCTTAATCTATGTTGGCAATT 2643
Qy 803 ThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAla 822
Db 2644 ACACATTTTACTCTTAGCAATTTGATCCCATCATTCGAATAGCAGGAGTTGTTGAAATG 2703
Qy 823 AlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAla 842
Db 2704 ATGTTGTCTGGCAAGCACTGAAAGATAGAAAGAGCTAGAAAGAGCTGGGAAGATTGT 2763
Qy 843 ThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGlu 862
Db 2764 ACAGAAGCCATCGAAACTTCGAACTGTTGTTTCTTGACTCGGAGCAGAAGTTTGA 2823
Qy 863 GlnMetTyrGluGluMetLeuGlnHisArgAsnThrSerLysLysAlaGlnIle 882
Db 2824 TACATGTATGCACAGATTTGCAAGTACCATACAGAAACTCTTGAGGAAGACACATC 2883
Qy 883 IleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaGlyPhe 902
Db 2884 TTCGGGCTCTCAATTTCTATCATCCAGCAATGATGATATTTTCTATATGCTGGCTTTC 2943
Qy 903 ArgPheGlyValaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleVal 922
Db 2944 CGGTTTGGTCCCTACTTGTGGCAATGATGTTCAATGAACTTTTCAGGATGTTCTTTGGTA 3003
Qy 923 PheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGlu 942
Db 3004 TTTCTCAGCTATTGCTTTGTTGGTCCATGGCAGTGGGCGAGTTCAGTTTCATTTGCTCCTGAC 3063
Qy 943 TyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuLysLysProAsn 962
Db 3064 TATGCCAAAGCCAAAGTATCAGCAGCCACGTCATCATCATCATTTGAAAGAAAGCCCTCTG 3123
Qy 963 IleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPhe 982
Db 3124 ATTGACAGCTACAGCCCTCAGGCTCAGCCCAATACATTTGGAAGAAATGTGACATTT 3183
Qy 983 ArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSer 1002
Db 3184 AATGAGTCTGTCTCAACTATCCACTCGACACAGCATCCCGTGTCTCCAGGCGGTGAGC 3243
Qy 1003 LeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSer 1022
Db 3244 CTCAGGTTGAAGAAGGCCAGACGCTGGCCCTCGTAGGTAGCAGTGGCTGTGGGAAGAGC 3303
Qy 1023 ThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAsp 1042
Db 3304 ACAGTTGTTGAGTCTCAGAGCGCTTCTATGACCCCTTGGCTGGTTCAGTGTCTAATGAT 3363
Qy 1043 GlyValAspAlaLysGluLeuAsnValGlnTyrLeuArgSerGlnIleAlaIleValPro 1062
Db 3364 GGCAAGAGATAAAGCACCTGAATGTCCAGTGGCTCCGAGCACACCTGGGCGCTGCTCT 3423
Qy 1063 GlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSer 1082
Db 3424 CAGGAGCCCATCTGTTTGAATGCTGAGCATTTGCCGAAACATTTGCCCTATGGAGCAACAGC 3483
Qy 1083 ArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPhe 1102
Db 3484 CGGTCGTATCATGAAGAGATTATGCGAGGCGCCAGGAGGCCAACATACACCTTC 3543

QY 1103 IleGluGlyLeuProGluLysTyrAenThrGlnValGlyLeuLysGlyAlaGlnLeuSer 1122
 DB 3544 ATCGAGACACTCCCTGAGAAATACACACAGAGTAGGAGACAAAGAACCCAGCTCTCT 3603
 QY 1123 GlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeu 1142
 DB 3604 GGTGGCCAGAAACAGCGCATTCGCATAGCTCGCGCTCTTGTAGACAGCCCTCATATTTG 3663
 QY 1143 LeuLeuAspGluAlaThrSerAlaLeuAspAsnSerGluLysValValGlnHisAla 1162
 DB 3664 CTTTGGATGAAGCTACATCAGCTCTGGATACAGAAAGTGTTCGAAGGCC 3723
 QY 1163 LeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIle 1182
 DB 3724 CTGGACAAACCCAGAGAGGCCGACCTGATGTGATCGCCACCGCTTGTCCACCATC 3783
 QY 1183 GlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHis 1202
 DB 3784 CAGAATGCAGATTTAATAGTGGTGTTCAGAAATGGCAAGTCAAGGAGCATGGCACACAT 3843
 QY 1203 GlnGluLeuLeuArgAsnArgAspIleTyrPhelLysLeuValAsnAlaGlnSer 1220
 DB 3844 CAACAGCTGCTGCTCAGAAAGCACTATTTTCCATGATCAGTGTCCAGGCT 3897
 RESULT 10
 AAF86127
 ID AAF86127 standard; cDNA; 4186 BP.
 XX AAF86127;
 AC AAF86127;
 XX AAF86127;
 DT 25-JUN-2001 (first entry)
 XX Cynomologous monkey P-glycoprotein cDNA.
 DE Cynomologous monkey P-glycoprotein cDNA.
 XX Cynomologous monkey; P-glycoprotein; PGP; multidrug transporter; MDR1;
 KW efflux pump; ss.
 XX Macaca fascicularis.
 OS
 XX
 PH Key Location/Qualifiers
 FT CDS 100..3942
 FT /*tag= a
 FT /product= "PGP"
 FT /note= "P-glycoprotein"
 XX
 PN WO200123565-A1.
 XX
 PD 05-APR-2001.
 XX
 XX 28-SEP-2000; 2000WO-US26592.
 XX
 XX 28-SEP-1999; 99US-0156921.
 PR 12-OCT-1999; 99US-0158818.
 XX
 XX (GENT-) GENTEST CORP.
 PA
 XX
 XX Stocker PJ, Steimel-Crespi DT, Crespi CL;
 XX
 DR WPI; 2001-316136/33.
 DR P-PSDB; AAB81064.
 DR
 XX Novel isolated nucleic acid encoding cynomologous monkey P-glycoprotein
 PT (PGP) and homologous PGP polypeptides are useful for predicting
 PT bioavailability of compound and increasing PGP transporter activity in
 PT cell -
 XX
 XX Example 1; Page 51-57; 84pp; English.
 PS
 XX
 XX This invention relates to a polynucleotide sequence encoding a
 CC cynomologous monkey P-glycoprotein (PGP), and an allelic variant of the
 CC PGP protein. PGP, also known as multidrug transporter, MDR1 is a member
 CC of the ABC transporter superfamily. The enzyme serves as an efflux pump
 CC exporting small molecules across the cell membrane. The invention

CC includes a cynomologous monkey (Macaca fascicularis) PGP coding sequence
 CC and protein, and also that of an allelic variant. The PGP polynucleotide
 CC sequence is useful for increasing PGP transporter activity in a cell.
 CC Antisense sequences of the cDNA are useful for inhibiting PGP transport
 CC activity in a mammalian cell. They may also be used for increasing the
 CC bioavailability of a drug. The present sequence represents cDNA encoding
 CC cynomologous monkey P-glycoprotein.
 XX
 SQ Sequence 4186 BP; 1226 A; 799 C; 1039 G; 1122 T; 0 other;
 Alignment Scores:
 Pred. No.: 0 Length: 4186
 Score: 3524.50 Matches: 679
 Percent Similarity: 73.75% Conservative: 248
 Best Local Similarity: 54.02% Mismatches: 261
 Query Match: 57.23% Indels: 69
 DB: 22 Gaps: 11
 US-09-873-409-5 (1-1222) x AAF86127 (1-4186)
 QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
 DB 250 ATGGTGGTGGAACTTGGCTGCCATCATCCATGGAGCTGGACTTCCTCTCATGATGCTG 309
 QY 21 ValLeuGlyGluMetSerAspAsnLeuIleSer-----GlyCys 33
 DB 310 GTGTTTGGAGACATGACGATACCTTTTGCAAATGCGAGGAAATTTAGGAGATTTAGGAGCT 369
 QY 34 LeuValGlnThrAsnThrTyrSer-----41
 DB 370 CTGTTG---ACTAATAGCAGTAAATATCATCATGTCGCCCTCATGATCTGGAGGAA 426
 QY 42 ---PhePheArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePheGly 60
 DB 427 GATATCACCAGGTATGCTTATTATACAGTGGAAATGGTGGGGTGGTGGTGGTGGTGGT 486
 QY 61 TyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLys 80
 DB 487 TACATTGAGTTTCATTTTGGTGGCTGGCAGCTGGGAAGACAAATACACAAATTAGAAAA 546
 QY 81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGly 100
 DB 547 CAGTTTTTTCATGCTATAATGCGACAGGATAGGCTGGTTTATGTGACGAGTGTGGG 606
 QY 101 GluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLys 119
 DB 607 GAGCTTAACACCGCGCTTACAGATGATGTCTCAAGATTAATGAGGAATTTGTGACAAA 666
 QY 120 IleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuVal 139
 DB 667 ATTGGAATGTTCTTTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGGATTTACA 726
 QY 140 LysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAla 159
 DB 727 CGTGGTTTGAAGCTAAACCTTGTGATTTTGGCCATCAGTCTCTTCTTGACTGTCTAGCT 786
 QY 160 AlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLys 179
 DB 787 GCAGCTGGGCAAGATACCTGCTTCTATTACTGATATAAGAACTCTTAGCTTATGCAAAA 846
 QY 180 AlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAla 199
 DB 847 GCTGGAGCAGTAGCTGAAGAGGCTTGGCAGCAATTAGAAGCTGTGATGATGATTTGGAGGA 906
 QY 200 GlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTyrPheTyr 219
 DB 907 CAAAAGAAAGAACTCGAA-----924
 QY 220 PheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLys 239
 DB 925 -----AGGTACACAAATAATTTAGAA 945
 QY 240 AspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaVal 259

Db 946 GAAGCTAAAGAAATTGGGATAAGAAAGCTATTACAGCCAATATTCTATAGGTGCTGCT 1005
Qy 260 TyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrGlyThrSerLeuLeuLeu 279
Db 1006 TTCCTGCTTATCATGCACTTATGCTCTGGCTTCTGATGGAGCACTTGGTCTC 1065
Qy 280 AsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHis 299
Db 1066 TCAAAAGGAA-----TATTCTATTGGACAAGTACTCACTGTATTCTTTCTGTATTAA 1119
Qy 300 SerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGly 319
Db 1120 GGGGCTTTTAGTGTGGACAGGCATCTCCAAGCATTTGAAGCATTTGCAATGCAAGAGGA 1179
Qy 320 AlaAlaPheHisIlePheGlnValIleAspLysPheProSerIleAspAsnPheSerThr 339
Db 1180 GCAGCTTTTGAATCTTCAAGATATTGATTAATAGCCAACTATTGACAGCTATTGCAAG 1239
Qy 340 AlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsn 359
Db 1240 AGTGGGCACAAACCATTAATTAAGGAAATTTGGAATTCAGAAATGTTCACTTCAGT 1299
Qy 360 TyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGly 379
Db 1300 TACCCATCTCGAAAGAAAGTAAAGATCTTGAAGGGCTCGAACCTGAAGGTGCAGAGTGG 1359
Qy 380 GluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeu 399
Db 1360 CAGACGGTGGCCCTGGTGTGGAACACGGCTGTGGAGAGACAAACGGTCCAGCTGATG 1419
Qy 400 GlnArgLeuTyrAspProAspGlyPheIleMetValAspGluAsnAspIleArgAla 419
Db 1420 CAGAGCTTTATGACCCACAGAGGCGCATGGTCAGTGTGATGACAGGATATTAGGACC 1479
Qy 420 LeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPhe 439
Db 1480 ATAAACGTAAAGTGTCTACGGGAAATCATCGGTGTGGTGTGAGTCAGGAACCTGTATTGTT 1539
Qy 440 GlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMet 459
Db 1540 GCCACCATGATGCTGAAGACCATTCCTATGCTATGCTGTGGAAGATGTCACCATGGATGAT 1599
Qy 460 GluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPhe 479
Db 1600 GAGAAAGCTGTCAAGAGGACCAATGCTATGACTTTATCATGAACCTGCTCAGAAATTT 1659
Qy 480 AsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAla 499
Db 1660 GACACCTGCTGGAGAGAGGGGCCAGCTGAGTGTGGTGGGAGAGGAGGATCGCC 1719
Qy 500 IleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAla 519
Db 1720 ATTGCAAGTGCCTGCTGTGCAACCCCAAGATCCTCTGCTGGACGAGGCGACGTCAGCC 1779
Qy 520 LeuAspSerGluSerLysSerAlaValGlnAlaLeuGluLysAlaSerLysGlyArg 539
Db 1780 TTGGACACAGAAAGTGAAGCAGTGGTTCAGTGTCTGATTAAGGCCAGAAAGGTCCG 1839
Qy 540 ThrThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThr 559
Db 1840 ACCACCATTTGATAGTCACTGCTTTGCTACGGTTGCTAATGCCGACGTCATCGTGGT 1899
Qy 560 LeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGly 579
Db 1900 TTCGATGATGGACTCATTTGGAAGAAAGGAAATCATCATGAGCTCATGAAGAGAAAGGC 1959
Qy 580 LeuTyrTyrSerLeuValMetSerGlnAspIle-----LysLys 592
Db 1960 ATTTACTTCAAACTTGTCAAAATGTCAGACAGCAGGAAATGAAATGAAATGAAATGCA 2019
Qy 593 AlaAspGlu-----GlnMetGluSerMetThrTyrSerThrGluArgLysThrAsn 609
Db 593 AlaAspGlu-----GlnMetGluSerMetThrTyrSerThrGluArgLysThrAsn 609

Db 2020 GCTGATGATCCAAAAGTGAAATTGATACCTTGGAAATGCTCTTCACATGATTCAGATCC 2079
Qy 610 SerLeu-----ProLeuHisSerValLysSerIleLysSerAspPheIle 624
Db 2080 AGTCTAATAAGAAAAAGATCCACTCGTAGGAGTGTCCGTGGATCATCAAGGCCAA----- 2133
Qy 625 AspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeu 644
Db 2134 GACAAAAAGCTTAGTACCAAGAGGCTCTGGATGAAGATATACCTCCAGATTTCTTTTGG 2193
Qy 645 LysIleLeuLysLeuAsnLysProGluTyrProPheValValLeuGlyThrLeuAlaSer 664
Db 2194 AGGATTAAGCACTAAATTAACAGTGGCTTATTTGTTGTTGGTATTGTTGCTG 2253
Qy 665 ValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMet 684
Db 2254 ATTATAAATGGAGGTCTGCAACACGAGCATTTGTCAGTAATATTTTCAAGATTTATAGG 2313
Qy 685 Phe---GlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIle 703
Db 2314 TTTACAAGAAATGATGATGCCGAAACAAACACAGCAATAGTAACCTGTTTTCACATTTG 2373
Qy 704 PheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGly 723
Db 2374 TTTCTAGTCTTGGAAATGTTTCTTTTATATATTTTCTTTCCTTCAGGCTTCACATTTG 2433
Qy 724 ArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyr 743
Db 2434 AAAGCTGGAGAGATCTCACCAAGCGGCTCCGATACATGTTTTCGATCCATGCTCAGA 2493
Qy 744 GlnAspIleAlaTyrPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeu 763
Db 2494 CAGGATGTGAGTGGTGTGATGACCTTAAACACCACTGGAGCATTTGACTACCAAGCTC 2553
Qy 764 AlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGln 783
Db 2554 GCCATGATGCTGCTCAAGTTAAGGGCTATAGTTCAGGCTGCTATATATATATATAT 2613
Qy 784 AsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyrGluMetThr 803
Db 2614 AATATAGCAAACTTTGGGACAGGAATAATATATCTTAATCTATGTTGGCACTGACA 2673
Qy 804 PheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAla 823
Db 2674 CTGTACTCTTAGCAATTTGATCCCATCATTTGCAATAGCAGGAGTGTGGAATGAAATG 2733
Qy 824 MetThrGlyPheAlaAsnLysAspLysGlnLeuLysHisAlaGlyLysIleAlaThr 843
Db 2734 TTGCTGGACAAGCACTGAAGATGAAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACT 2793
Qy 844 GluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGln 863
Db 2794 GAAGCAATAGAAAACCTTCGAACTGTTGTTTCTTGACTCAGGAGCAGAAGTTTGAACAT 2853
Qy 864 MetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIle 883
Db 2854 ATGTATGATCAGAGTTTGAGGTACCATACAGAAAACCTTTTGAGGAAAGCACACATCTTT 2913
Qy 884 GlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArg 903
Db 2914 GGAATCACGCTTTCTTCACGCGCAATGATGATTTTTTCTATGCTGATGATTTTCCCG 2973
Qy 904 PheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPhe 923
Db 2974 TTTGGAGCTTACTTGTGGCACATAGTCTCATGAGCTTTGAGGATGTTCTGTAGTATTT 3033
Qy 924 ThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyr 943
Db 3034 TCAGCTGTTGTTGTCATGCGCATGGCGTGGGCAAGTCAGTTTCATTTGCTCTGACTAT 3093
Qy 944 SerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysProAsnIle 963
Db 3094 GCCAAAGCCAAAAGTATCAGCAGCCCATCATCATGATCATTTGAAAAAACCCTTTGATT 3153

QY 964 AspSerArgSerGlnGluGlyLysProAspThrCysGluGlyAsnLeuGluPheArg 983
 DB 3154 GACAGCTACAGCAGAGAGCCCTAAGCCGAAACACATTGGAGGAATGTCACATTAAT 3213
 QY 984 GluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeu 1003
 DB 3214 GAAGTTGTATTCAACTATCCACCGACTGGACATCCAGTGTTCAGGGCTGAGCCTG 3273
 QY 1004 SerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThr 1023
 DB 3274 GAAGTGAAGAGGCGCAGACGCTGGCCCTGTGGCGGAGAGTGGCTGTGGGAAGACAGC 3333
 QY 1024 SerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlnValLeuPheAspGly 1043
 DB 3334 GTGGTCAGCTCTCTGGAGCGGTTCTATGACCCCTTGGCGGGAGAGTGTCTGTGACGGC 3393
 QY 1044 ValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGln 1063
 DB 3394 AAAGAAATAAAGCAACTGAATGTTCAAGTGCTCCGAGCACACCTGGGCATCGTGTCCAG 3453
 QY 1064 GluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArg 1083
 DB 3454 GAGCCCATCTCTGTGACTCGACCATTTAGTGAACATTCGCTATGGAGCAACAGCCCGG 3513
 QY 1084 ValValProLeuAspGluIleLysGluAlaAsnAlaAlaAsnIleHisSerPheIle 1103
 DB 3514 GTGGTGTACAGAGAGATCGTAGGGCAGCCAGGAGGCCAATATACAGCCTTCATC 3573
 QY 1104 GluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGly 1123
 DB 3574 GAGTCACTGCCTTAATAATATAGCACCAGAGTAGGAGACAAGAACTCAGCTCTCTGGT 3633
 QY 1124 GlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuGlnLysProLysIleLeuLeu 1143
 DB 3634 GGCAGAAACACAGCACTTGGCATAGCTGTCGCTCTGTAGACGCTCATATTTTGGCT 3693
 QY 1144 LeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeu 1163
 DB 3694 TTGATGAGCCACATCAGCTCTGGATACAGAAAGTGAAGAGTTGTCAGAGAGCCCTG 3753
 QY 1164 AspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGln 1183
 DB 3754 GACAAAGCCAGAGAGGCGCTACCTGATTTGTGATTGCTCACCGCTGTCCACCATCCAG 3813
 QY 1184 AsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGln 1203
 DB 3814 AATGCAGACTTAATAGTGGTGTTCAGAAATGGCAGAGTCAAGGAGACGCGCACATCAG 3873
 QY 1204 GluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAlaGlnSer 1220
 DB 3874 CAGCTGTGGCAGAGAGGCACTATTTTCAATGTGTAGTCCAGGCT 3924
 RESULT 11
 ID AA249333
 XX AA249333 standard; cDNA; 3860 BP.
 AC AA249333;
 XX
 DT 14-MAR-2000 (first entry)
 XX Human G185V mutant multidrug resistance-1 (MDR-1) cDNA.
 DE
 XX
 KW Multidrug resistance; MDR-1; P-glycoprotein;
 KW transmembrane efflux pump; haematopoietic stem cell; transduction;
 KW bone marrow transplantation; chemotherapy; radiation therapy; cancer;
 KW gene therapy; gene replacement; genetic defect; thalassaemia;
 KW Gaucher's disease; sickle cell anaemia; leukaemia; ex vivo expansion;
 KW cytokine; mutant; ds.
 OS
 OS Synthetic.
 OS Homo sapiens.
 XX

PH Key Location/Qualifiers
 FT CDS 1..3843
 FT /*tag= a
 FT /product= "Human G185V mutant MDR-1 protein"
 XX
 PN WO9961589-A2.
 XX
 PD 02-DEC-1999.
 XX
 PE 27-MAY-1999; 99WO-US11825.
 XX
 PR 28-MAY-1998; 98US-0086988.
 XX
 PA (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
 PI Sorrentino B, Bunting K;
 DR WPI; 2000-072615/06.
 DR P-PSDB; AAY58187.
 XX
 PT Ex vivo expansion of hematopoietic stem cells transduced with a
 PT sequence encoding human multidrug resistance-1, used for bone marrow
 PT transplantation -
 XX
 PS Example 1; Page 79-82; 113pp; English.
 CC This sequence represents cDNA encoding human G185V mutant multidrug
 CC resistance protein MDR-1, where the Gly residue at position 185
 CC of the wild-type protein (AAY58186) is replaced by Val. MDR-1 is a
 CC transmembrane efflux pump, responsible for the export of drugs from
 CC cells, particularly cancer cells. The wild-type MDR-1 shows increased
 CC resistance to etoposide and decreased resistance to vinca alkaloids
 CC compared with the G185V mutant. The invention relates to transducing
 CC haematopoietic stem cells with nucleic acid encoding an MDR protein
 CC and culturing the modified cells. The modified haematopoietic stem
 CC cells are useful in bone marrow transplantation (to reconstitute
 CC haematopoietic systems in patients who have undergone chemotherapy or
 CC radiation therapy) and in ex vivo gene therapy of genetic defects in
 CC cells derived from haematopoietic stem cells, e.g., thalassaemia,
 CC Gaucher's disease, sickle cell anaemia or leukaemia. The modified
 CC cells can also be used to identify factors involved in regulating
 CC proliferation and differentiation in haematopoietic stem cells.
 CC Haematopoietic stem cells that express MDR-1 will be protected against
 CC chemotherapeutic agents, so can be engrafted while the patient is
 CC undergoing chemotherapy. Expansion of (rare) haematopoietic stem cells
 CC provides sufficient cells to permit standard biochemical analysis.
 CC Overexpression of MDR-1 allows cytokine-driven expansion of
 CC haematopoietic stem cells by at least 10-fold compared with a maximum
 CC of 4-fold in known procedures.
 XX
 SQ Sequence 3860 BP; 1135 A; 746 C; 956 G; 1023 T; 0 other;
 Alignment Scores:
 Pred. No.: 0 Length: 3860
 Score: 3521.50 Matches: 681
 Percent Similarity: 73.92% Conservative: 246
 Best Local Similarity: 54.31% Mismatches: 264
 Query Match: 57.18% Indels: 63
 DB: 21 Gaps: 10
 US-09-873-409-5 (1-1222) x AA249333 (1-3860)
 QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
 DB 151 ATGGTGGTGGAAACTTGGCTGCCATCATCCATGGGGCTGGACTTCTCTCATGATGCTG 210
 QY 21 ValIleuGlyGluMetSerAsp-----AsnLeuIleSer 31
 DB 211 GTCTTTGGAGAAATGACAGATATCTTGCAAATGCAGGAAATTTAGAGATCTGTGATGCA 270
 QY 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
 DB 271 AACATCACTAATAGAAAGTGATATCAATGATACAGGGGTCTTTCATGATCTGGAGGAGAC 330

QY 44 -----ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
DB 331 ATGCACAGATATGCTATTATTACAGTGAATTTGGTCTGGGGTGTCTGTTCTGCTTAC 390
QY 62 IleGlnIleSerLeuTyrIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
DB 391 ATTCAGGTTTCATTTTGGTCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 450
QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101
DB 451 TTTTTCATGCTATAATGCGACAGGAGATAGCTGGTTGATGTCACCATGTTGGGAG 510
QY 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
DB 511 CTTAAACCCGACTTACAGATGATGCTCTAAGATTAAATGAAGTTATTGCTGACAAAATT 570
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
DB 571 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGGATTTACAGCT 630
QY 141 GlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
DB 631 GGTGGAGCTAACCTTGTGATTGGCCATCAGTCTCTGTTGTCCTGGACTGTCAGCTGCT 690
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
DB 691 GTCGGCAAGATACACTCTTCATTTCTGATAAAGAACTCTTAGCGTATGCMAAGCT 750
QY 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
DB 751 GGAGCAGTAGCTGAAGAGCTTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 810
QY 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTyrPheTyrPhe 220
DB 811 AGAAAGAACTTGAA----- 825
QY 221 ProGlnTyrLeuLeuSerCysValLeu**PheValArgTyrThrGlnAsnLeuLysAsp 240
DB 826 -----AGGTACACAAACAAATTTAGAGAA 849
QY 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
DB 850 GCTAAAGAAATTTGGGATAAAGAAAGCTATTACAGCCAAATATTTCTATAGTGTGCTTTC 909
QY 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsn 280
DB 910 CTGCTGATCTATGCTATGCTCTGGCCCTCTGGTATGGGACCACCTTGGTCTCTCA 969
QY 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
DB 970 GGGGA-----TATTCTATTGGACAAGTACTCCTGTTCTTTCTGTTAATTGGG 1023
QY 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
DB 1024 GCTTTTAGTGTGGACAGGCATCTCCAAAGCATTTGAAGCATTTTCCAAATGCAAGAGGACA 1083
QY 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
DB 1084 GCTTATGAAATCTTCAAGATAATTGATAATAAGCCAAAGTATTGACAGCTATTCCGAAGAGT 1143
QY 341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
DB 1144 GGCACAAACAGATATAATTAGGGAATTTGGAATTCGAATGTTCACTTCAGTTAC 1203
QY 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
DB 1204 CCATCTCGAAAAGAGTAAAGATCTTGAAGGCTTGAACCTGAAGTGCAGAGTGGGAC 1263
QY 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
DB 1264 ACGGTGGCCCTGTTGGAAACAGTGGCTGTGGGAAGAGACAAACAGTCCAGCTGATGCAG 1323

QY 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
DB 1324 AGGCTCTATGACCCCCACAGAGGGGATGTCAGTGTGTGATGGACAGCATATTAGGCCATA 1383
QY 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
DB 1384 RATGTAAAGTTTCTACGGGAATCATTTGGTGGTGTGAGTCAGAACCTGTATTGTTGCC 1443
QY 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluMetGlu 460
DB 1444 ACCACGATAGCTGAAACATTCGCTATGCCCGTGAAATGTCCACATGATGAGATTGAG 1503
QY 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
DB 1504 AAAGCTGTCAAGGAACCAATGCCCTATGACTTTATCATGAATGCCCTCATAAATTTGAC 1563
QY 481 ThrLeuValGlyLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
DB 1564 ACCCTGGTTGGAGAGAGGGGCCAGTTGAGTGTGGCAGACAGAGGATCGCCATT 1623
QY 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
DB 1624 GCACGTGCCCTGGTTCGCAACCCCAAGATCCTCTGCTGGATGAGGCCACGTCAGCCTTG 1683
QY 521 AspSerGluSerLysSerAlaValGlnAlaLeuGluLysAlaSerLysGlyArgThr 540
DB 1684 GACACAGAAAGCGAGCAGATGTTCCAGGTGGCTCTGGATAAGGCCCAAAAGGTGGACC 1743
QY 541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560
DB 1744 ACCATTGATAGCTCATCGTTTGTCTACAGTTCGTAAATGCTGACGTCATCCTGCTGTTTC 1803
QY 561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580
DB 1804 GATGATGGAGTCATTGTGGAGAAAGAAATCATGTAACCTCATGAAAGAGAAAGCATT 1863
QY 581 TyrTyrSerLeuValMetSerGln-----AspIleLysLysAlaAsp 594
DB 1864 TACTTCAAACTTGTCCACATGCAGACAGCAGGAAATGAAGTTGAATTTAGAAAATGCAGCT 1923
QY 595 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis 614
DB 1924 GATGAATCCAAAAGTGAATTTGATGCTTGGAAATGCTCTTCAATGATTCAGATCCAGT 1983
QY 615 SerVal-----LysSerIleLys---SerAspPheIleAspLysAla 627
DB 1984 CTAATAAGAAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAGCCCAACAGACAGAAAG 2043
QY 628 GluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeu 647
DB 2044 CTTAGTACCAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCTCTTTTGGAGGATTATG 2103
QY 648 LysLeuAsnLysProGluTyrProPheValValLeuGlyThrLeuAlaSerValLeuAsn 667
DB 2104 ARGCTAAATTTAACTGAATGGCTTATTTTGTGTGGTATTTTGTGCCATTTATAAT 2163
QY 668 GlyThrValHisProValPheSerIlePheAlaLysIleIleThrMetPheGlyAsn 687
DB 2164 GGAGCCTGCACACGACATTTGCCAATAATATTTTCAAGATATTAGGGGTTTTCACAAGA 2223
QY 688 ---AsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIle 706
DB 2224 ATTGATGATCTCTGAAACAAACAGACAGATAAGTAACTGTTTCTTCTTCTTCTTCTAGCC 2283
QY 707 LeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGly 726
DB 2284 CTTGGAATTTATTCTTTTATTACATTTTCTTCAAGGTTTCCACATTTTGGCAAGCTGA 2343
QY 727 GluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIle 746
DB 2344 GAGATCTCCACCAAGCGCTCCGATACATGTTTCCGATACATGCTCAGACAGGATG 2403
QY 747 AlaTyrPheAspGluLysGluAsnSerThrGlyLeuThrThrIleLeuAlaIleAsp 766

Db 2404 AGTTGGTTTATGACCCATAAAGACACACTGGAGCATTTGACTACCAAGGCTCGCCATGAT 2463
Qy 767 llaaaglnleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThr 786
Db 2464 GCTGCTCAAGTTAAAGGGCTATAGGTTTCAGGCTTCTGTAATATACCCAGATATAGCA 2523
Qy 787 AsnMetGlyLeuSerValIleIleSerPheIleIleGlyTyrGluMetThrPheLeuIle 806
Db 2524 AATCTTGGGACAGGAATAATATATCTTCTCATCTATGTTGGCACTAAACACTGTTACTC 2583
Qy 807 LeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGly 826
Db 2584 TTAGCAATTTGACCATCAATGCAATAGCAGGAGTTGTTGAAATGAAATGTTGCTCTGGA 2643
Qy 827 PheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeu 846
Db 2644 CAAGCACTGAAGATAGAAAGAACTAGAGAGGTCGCGGAGATCGCTACTGAAGCAATA 2703
Qy 847 GluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGlu 866
Db 2704 GAAACCTTCGAACCGTTGTTCTTCTGACTCAGGAGAGAGTTTGACATATATGATGCT 2763
Qy 867 GluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCys 886
Db 2764 CAGAGTTTGCAGGTACCATACAGAACTCTTTGAGGAAAGCACACATCTTTGGAATTAACA 2823
Qy 887 TyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAla 906
Db 2824 TTTTCTTCCACCCAGGAATGATGATATTTTCTCATCTGAGTGTTCCTGCTGGAGCC 2883
Qy 907 TyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIle 926
Db 2884 TACTTGGGCACATAACTCATGAGTTTGAGGATGTTCTGTAGTATTTTCAGCTGTT 2943
Qy 927 AlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAla 946
Db 2944 GTCCTTGGTGCCATGGCGTGGGCAAGTCAGTTTCATTTGCTGCTGACTATGCCAAAGCC 3003
Qy 947 LysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArg 966
Db 3004 AAAATATCAGCAGCCCATCATCATCATGATCAATGAAAAACCCCTTTGATTTGACAGCTAC 3063
Qy 967 SerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSer 986
Db 3064 AGCAGGAAGCCCTAATGCCGAACACATTTGGAAGAAATGTCACATTTGGTGAAGTTGTA 3123
Qy 987 PhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGlu 1006
Db 3124 TTCAACTATCCACCCAGCCGACATCCAGTGTCTCAGGAGCTGAGCCTGGAGGTGAAG 3183
Qy 1007 ArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGln 1026
Db 3184 AAGGCCACAGCGCTGCTGCTGGCAGCAGTGGCTGTGGGAAGAGCACAGTGTGCTCCAG 3243
Qy 1027 LeuLeuGlnArgLeuTyrAspProValGlnGlyValLeuPheAspGlyValAspAla 1046
Db 3244 CTCCTGGAGCGGTTCTACGACCCCTTGGCAGGAAGTGTCTGCTGTGATGGCAAGAAATA 3303
Qy 1047 LysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProVal 1066
Db 3304 AAGCGACTGAATGTTAGTGGCTCCGAGCACACCTGGGSCATCGTGTCCAGGAGCCCATC 3363
Qy 1067 LeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValPro 1086
Db 3364 CTGTTTGACTCAGCATTTGCTGAGAACATTCCTATGAGAACACAGCCGGGTGTGTCA 3423
Qy 1087 LeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeu 1106
Db 3424 CAGGAAGAGATCGTGAGCGGACGAAAGAGGACCAATACATGCTTCATCGAGTCACTG 3483
Qy 1107 ProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlnLys 1126

Db 3484 CCTAATAATATATAGCACTAAAGTAGAGACAAAGGAAGTCTCTCTGGTGCCAGAAA 3543
Qy 1127 GlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuAspGlu 1146
Db 3544 CAACGCAATGGCCATAGCTGCTGCCCTTTGTAGACAGCCTCATATTTTGGTATGAA 3603
Qy 1147 AlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAla 1166
Db 3604 GCCACCTCAGCTCTGATACAGAAAGTGAAGAGTTGTCCAAAGAACCTGGCAAGGCC 3663
Qy 1167 ArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAsp 1186
Db 3664 AGAAGAGCCGACCTGCAATGTGATTTGCTACCCGCTGTCCACCATCCAGATGGAGAC 3723
Qy 1187 LeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeu 1206
Db 3724 TTATATAGTGGTGTTCAGATGGCAGAGTCAAGGACATGGCAGCATCAGCAGCTGCTG 3783
Qy 1207 ArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
Db 3784 GCACAGAAAGGCATCTATTTTCAATGGTCAGTGTCCAGGCT 3825
RESULT 12
ABA94366
ID ABA94366 standard; DNA; 3860 BP.
XX
AC ABA94366;
XX
DT 26-MAR-2002 (first entry)
XX
DE Human BCRP DNA related seq Id No. 3.
XX
KW Stem cell; ATP transport protein; ATP-binding cassette; antiparkinsonian;
KW hepatotrophic; neurodegenerative; cytostatic; antianemic; muscular; BCRP;
KW cardiant; gene therapy; ds.
XX
OS Homo sapiens.
FH
FH Key Location/Qualifiers
FT CDS 1..3843
FT /tag= a
XX
PN WO200192877-A2.
XX
XX 06-DEC-2001.
XX
XX 30-MAY-2001; 2001WO-US17459.
XX
XX 31-MAY-2000; 2000US-0584586.
PR 29-MAY-2001; 2001US-086866.
XX
XX (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX
XX Sorrentino B, Schuetz J;
PI
XX
XX WPI; 2002-114368/15.
DR P-PSDB; ABB07267.
XX
XX Identifying a stem cell, for treating e.g., muscular dystrophy,
PT myocardial infarction, Parkinson's disease, or neurodegenerative
PT disorders, comprises detecting the expression of an ATP transport
PT protein (BCRP) by a cell -
XX
XX Disclosure; Page 59-60; 87pp; English.
XX
XX The invention provides a method of identifying and/or isolating a stem
CC cell that involves detecting the expression of an ATP transport protein
CC containing a conserved ATP-binding cassette (BCRP) by a cell in a sample
CC comprising stem cells. The isolated stem cells may be used in the
CC treatment of diseases such as muscular dystrophy, degenerative liver
CC disorder, myocardial infarction, Parkinson's disease, degenerative
CC disorders of the brain, and for tissue regeneration or replacement.
CC Haematopoietic cells can be used in bone marrow transplants (e.g., for

CC treatment of leukemia) and for ex vivo gene therapy for treating blood
 CC diseases such as sickle cell anemia and thalassemia. The stem cells can
 CC also be used as cell targets in gene therapy protocols. The present
 CC sequence represents a sequence related to the BCRP for which no relevant
 CC information has been provided in the specification.

XX
 SQ Sequence 3860 BP; 1135 A; 746 C; 956 G; 1023 T; 0 other;

Alignment Scores:

Pred. No.: 0 Length: 3860
 Score: 3521.50 Matches: 681
 Percent Similarity: 73.92% Conservative: 246
 Best Local Similarity: 54.31% Mismatches: 264
 Query Match: 57.18% Indels: 63
 DB: 24 Gaps: 10

US-09-873-409-5 (1-1222) x ABA94366 (1-3860)

QY 1 MetIleuGlyIleuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
 DB 151 ATGGTGGTGGAACTTTGGCTGCATCATCCATGGGGCTGGACTTCTCTCATGATGCTG 210
 QY 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
 DB 211 GTGTTGGAGAAATGACAGATATCTTGGCAATGCGAGAAATTTAGAAGATCTGATGTCA 270
 QY 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
 DB 271 AACATCACTAATAGAGTATATCAATGATACAGGGTCTTTCATGATCTGGAGGAAGAC 330
 QY 44 -----ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaIlePheGlyTyr 61
 DB 331 ATGACCAGATATGCTTATTATACAGTGGAAATTTGGTGGTGGTGGTGGTGGTGGTGGT 390
 QY 62 IleGlnIleSerLeuThrIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
 DB 391 ATTCAAGTTTCATTTGGTGGCTGGCAGCTGGAGCAATATACACAAATTTAGAAACAG 450
 QY 82 PhePheIleSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101
 DB 451 TTTTTCATGCTATATATGCGACAGGAGATAGCTGGTTTATGTCACGATGTTGGGGAG 510
 QY 102 LeuAsnThrArgMetThr-----AspIleAspIleSerAspGlyIleGlyAspLysIle 120
 DB 511 CTTAACACCGCACTTACAGATGATGCTCTAAGATTAATGAAGTTATTTGGTCAGAAAT 570
 QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
 DB 571 GGAATGTCTCTTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGGATTTACACGT 630
 QY 141 GlyTyrPheLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
 DB 631 GGTGGAGACTAACCTTTGATTTTGGCCATCAGTCTGCTTCTTGGACTGTCAGCTGCT 690
 QY 161 AlaCysSerArgMetValIleThrLeuSerThrSerLysGluLeuSerAlaTyrSerLysAla 180
 DB 691 GTCTGGCAAGATGACTATCTTCAATCTGATAAAGAACTCTTAGCGTATGCAAAAGCT 750
 QY 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
 DB 751 GGAGCAGTAGCTGAAGAGGCTTTGGCAGCAATTTAGAACTGTGATTCATTTGGAGACAA 810
 QY 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTyrPheTyrPhe 220
 DB 811 AAGAAGAACTTGA----- 825
 QY 221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
 DB 826 -----AGGTACAACAAAAATTTAGAGAA 849
 QY 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
 DB 850 GCTAAAAGAAATTTGGGTAAAGAAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTC 909

QY 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsn 280
 DB 910 CTGCTGATCATGATCATCTTATGCTCTGGCTTCTGATGAGCAACCTTGGTCTCTCA 969
 QY 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
 DB 970 GGGGAA-----TATCTATTTGGACMACTACTGATTTCTTTCTGATTAATTTGGG 1023
 QY 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
 DB 1024 GCTTTTAGTGTGGACAGCATCTCCAAGCATTTGAAGCATTTGCAAAATGCAAGAGAGCA 1083
 QY 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
 DB 1084 GCTTATGAAATCTTCAAGATAATTAAGCCCAAGTATTGACAGCTATTTCGAAGAGT 1143
 QY 341 GlyTyrLysProGluSerIleGlyThrValGluPheLysAsnValSerPheAsnTyr 360
 DB 1144 GGGCACAACACAGATTAATTAAGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 1203
 QY 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
 DB 1204 CCATCTCGAAAGAAAGTAAAGATCTTGAAGGCGCTGAACCTGAAGTGCAGAGTGGCAG 1263
 QY 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
 DB 1264 ACGTGGCTGCTGTTGGAAACAGTGGCTGTGGAGAGACACACAGCTCCAGCTGATGCAG 1323
 QY 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
 DB 1324 AGGCTCTATGACCCACAGAGGGATGGTCAGTGTGTGAGTGGACAGGATATTAGGACCAT 1383
 QY 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
 DB 1384 NATGTAAAGTTTCTACGGGAAATCATTTGTGTGTGAGTCAAGAACTGTATTGTTGCC 1443
 QY 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluMetGlu 460
 DB 1444 ACCACGATAGCTGAAACATTCGCTATGTCGCGTGAATAATGTCCACCATGGATGAGATTGAG 1503
 QY 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
 DB 1504 AAAGTGTCAAGGAAGCCCAATGCCATATGATCTTATCATGAAACTGCTCATATAATTGAC 1563
 QY 481 ThrLeuValGlyLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
 DB 1564 ACCCTGTTGGAGAGAGAGGGCCAGTTGAGTGTGGCAGAGCAGAGGATCGCCATT 1623
 QY 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
 DB 1624 GCACGTGCTGCTGTTGCAACCCCAAGATCTCTCTGATGAGGCCACGTCAGCCTTG 1683
 QY 521 AspSerGluSerLysSerAlaValGlnAlaLeuGluLysAlaSerLysGlyArgThr 540
 DB 1684 GACACAGAAGCGACAGCTGTTTCAGGTGGTCTCTGGATAAGGCCAGAAAGTCGGACC 1743
 QY 541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560
 DB 1744 ACATTTGATGATGCTCATGTTGCTACATGTCGTAATGCTGACGTCATCGCTGTTTC 1803
 QY 561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580
 DB 1804 GATGATGGAGTCATTGTGGAGAAAGGAATCATGATGAATCATGAAGAGAAAGCAATT 1863
 QY 581 TyrTyrSerLeuValMetSerGln-----AspIleLysLysAlaAsp 594
 DB 1864 TACTTCAAACTGTCTCAATTCAGACAGCAGGAGAAATGAAGTTGAATTAGAAAATGCACT 1923
 QY 595 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis 614
 DB 1924 GATGAATCCAAAAGTGAATTTGATGCTTGGAAATGCTCTTCAAATGATTTCAAGATCCAGT 1983


```
QY 615 SerVal-----LysSerIleLys---SerAspPheIleAspLysAla 627
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 1984 CTAATAGAAAAGATCAACTCGTAGAGAGTCCGTGGATCAAGCCCAAGACAGAAG 2043
QY 628 GluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLysIleLeu 647
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 2044 CTTAGTACCAAGAGGCTCTGGATGAAGATATACCTCCAGTTTCCTTTTGGAGGATATG 2103
QY 648 LysLeuAsnLysProGluTyrProPheValValLeuGlyThrLeuAlaSerValLeuAsn 667
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 2104 ARGCTAAATTAACGTAATGGCTTATTTTGTGTGGTATTTTGTGGCATTAATAAT 2163
QY 668 GlyThrValHisProValPheSerIlePheAlaLysIleIleThrMetPheGlyAsn 687
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 2164 GGAGGCTGCAACACGATTTTGAATAATATTTTCAAGATATAGGGTTTTACAAGA 2223
QY 688 ---AsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIle 706
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 2224 ATTGATGATCCTGAAACAAAACACAGAGATAGTAACCTTGTTCACATATTGTTCTAGCC 2283
QY 707 LeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGly 726
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 2284 CTTGGAAATTAATCTTTTATTACATTTTCTTCAAGGTTTCACATTTGGCAAGCTGGA 2343
QY 727 GluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIle 746
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 2344 GAGATCCTCAACAGCGCTCCGATACATCGTTTTCGATCCATGCTCAGACAGGATGTG 2403
QY 747 AlaTyrPheAspGluLysGluAsnSerThrGlyLeuThrThrIleLeuAlaIleAsp 766
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 2404 AGTTGGTGTGATGACCCCTAAACACACCTGGAGCATTTGACTACCAAGCTCGCCAAATGAT 2463
QY 767 IleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThr 786
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 2464 GCTGCTCAAGTTAAAGGGCTATAGGTTCAGGCTTCTGTAATTACCCAGAATATAGCA 2523
QY 787 AsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyrGluMetThrPheLeuIle 806
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 2524 ANCTTGGGACAGGAATAATATATCTCTCATCTATGTTGGCAACTAACACTGTTTACTC 2583
QY 807 LeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGly 826
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 2584 TTAGCAATTTGATCCCATCAITGCAATAGCAGGAGTTGTTGAAATGAAATGTTGTCTGGA 2643
QY 827 PheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeu 846
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 2644 CAAGCACTGAAAGATAAGAAAGAACTAGAGGTCTGGAGAGATCGCTACTGGAAGCAATA 2703
QY 847 GluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGlu 866
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 2704 GAAAACTTCCGAACCGTGTGTTCTTCTGACTCAGGAGCAGAAGTTTGAACATATGTATGCT 2763
QY 867 GluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCys 886
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 2764 CAGAGTTTGGCAGGTACCATACAGAAACTCTTTGAGGAAGACACACATCTTTTGGAAATTACA 2823
QY 887 TyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAla 906
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 2824 TTTTCTTCCACCGCAATGATGATATTTTCTATGCTGGATGTTTCCGGTTTGGAGCC 2883
QY 907 TyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIle 926
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 2884 TACTTGTGGCACATAAATCATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTT 2943
QY 927 AlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAla 946
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 2944 GTCCTTGTGTCATGGCGTGGGCAAGTCAGTTTCATTTGCTCTGCTGACTATGCCAAAGCC 3003
QY 947 LysSerGlyAlaAlaHisLeuPheAlaLeuLeuGlyLysProAsnIleAspSerArg 966
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 3004 AAAATATCAGAGCCCATCATCATGATCATTTGAAAAAACCCCTTTGATTGACAGCTAC 3063
QY 967 SerGlnGluGlyLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSer 986
```

```
Db 3064 AGCACGAAGGCGCTAATGCCGAACACATTTGGAAGGAAATGTACATTTGGTGAAGTTGTA 3123
QY 987 PhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGlu 1006
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 3124 TTCACTATCCCAACCCGACCGACATCCCGAGCTTTCAGAGGATGAGCTTGGAGGTGAAG 3183
QY 1007 ArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGln 1026
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 3184 AAGGGCCAGACGCTGCTCTGTGGGCGACAGTGGCTGTGGGAAGACACAGTGGTCCAG 3243
QY 1027 LeuLeuGlnArgLeuTyrAspProValGlnGlnValLeuPheAspGlyValAspAla 1046
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 3244 CTCCTGGAGCGGTTCACGACCCCTTGGCAGGGAAGTGTCTTGTATGCAAGAAATA 3303
QY 1047 LysGluLeuAsnValGlnTyrPheLeuArgSerGlnIleAlaIleValProGlnGluProVal 1066
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 3304 AAGCGACTGAATGTTCCAGTGGCTCCGAGCACACCTGGGATCGTGTCCAGGAGCCCATC 3363
QY 1067 LeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValPro 1086
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 3364 CTGTTTGACTGCGACGATTCGTGAGAACATTTGCCATATGGAGACAACAGCCGGTGTGTC 3423
QY 1087 LeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGlyGlyLeu 1106
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 3424 CAGGAAGAGATCGTGAGGCGCAGCAAGAGGCGCAACATACATCCCTTCATCGAGTCACTG 3483
QY 1107 ProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLys 1126
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 3484 CCTAATAATATAGCACTAAAGTAGGAGACAAGAACTCAGCTCTCTGGTGGCCAGAAA 3543
QY 1127 GlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuAspGlu 1146
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 3544 CAACGCAATGCCATAGCTCGTGCCCTTGTGTAGACAGCCCTCATATTTTGTGTTGATGAA 3603
QY 1147 AlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAla 1166
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 3604 GCCACGCTCAGCTCTGGATACAGAAAGTGAAGAGTGTGTCCAAGAAAGCCCTGGCAAGCC 3663
QY 1167 ArgThrGlyArgThrCysLeuValThrHisArgLeuSerAlaIleGlnAsnAlaAsp 1186
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 3664 AGAAGAGCGCGCACCTGCTGATTTGATTTGCTCACCCTGTCCACCATCCAGAAATGCAGAC 3723
QY 1187 LeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeu 1206
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 3724 TTAATAGTGTGTTTCAGAAATGSCAGATCAAGAGCATGCGCAGCATCAGCAGCTGCTG 3783
QY 1207 ArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 3784 GCAGAAAGGCATCTATTTTCAATGGTCAGTGTCCAGGCT 3825
```

RESULT 13

AAH57442
ID AAH57442 standard; cDNA; 4349 BP.

XX AAH57442;

DT 10-SEP-2001 (first entry)

XX Human intestine cell specific cDNA sequence SEQ ID NO:282.

XX Human; tissue specific; diagnosis; brain; heart; skeletal muscle;
KW lung; liver; uterus; ovary; stomach; intestine; kidney; pancreas; ss;
KW metabolic disease; developmental disease; cystic; immunomodulatory;
KW neuroprotective; gene therapy; cancer; immunopathology; neuropathology.

OS Homo sapiens.

XX WO200132927-A2.

XX 10-MAY-2001.

XX 02-NOV-2000; 2000WO-US30396.

XX 04-NOV-1999; 99US-0163508.
XX (INCY-) INCYTE GENOMICS INC.
XX Sornasse T, Seilhamer JJ, Watson GA;
XX WPI; 2001-291057/30.
XX New cell and tissue specific polynucleotides useful for diagnosis,
XX prognosis or monitoring of treatments for disorders where the gene is
XX associated with a cancer, immunopathology or neuropathology -
XX Claim 1; Page 207-208; 327pp; English.
XX AAH57161 to AAH57576 represent cell and tissue specific polynucleotide
XX sequences (I). (I) can have cytostatic, immunomodulatory and
XX neuroprotective activities, and can be used in gene therapy. (I) and
XX proteins (II) encoded by then are used in high throughput screening
XX assays to select DNA molecules, RNA molecules, peptide nucleic acids,
XX mimetics, peptides, proteins, agonists, antagonists, antibodies or
XX their fragments, immunoglobulins, inhibitors, drug compounds and
XX pharmaceutical agents. Expression of (I) in a sample indicates the
XX differentiation of embryonic stem cells into a tissue selected from
XX brain, heart, kidney, liver, lung, skeletal muscle or pancreatic
XX tissues. (I) and (II) are used to produce an expression profile that
XX defines a metabolic or developmental process, treatment, condition,
XX disease or disorder. The gene profile can be used for diagnosis,
XX prognosis or monitoring of treatments and for investigating a
XX predisposition to a disorder where the gene is associated with a
XX cancer, immunopathology or neuropathology.
XX SQ Sequence 4349 BP; 1304 A; 816 C; 1048 G; 1181 T; 0 other;

Alignment Scores:
Pred. No.: 0 Length: 4349
Score: 3521.50 Matches: 681
Percent Similarity: 73.9% Conservative: 246
Best Local Similarity: 54.31% Mismatches: 264
Query Match: 57.18% Indels: 63
DB: 22 Gaps: 10

US-09-873-409-5 (1-1222) x AAH57442 (1-4349)
QY 1 MetIleuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
DB 276 ATGTGGTGGAACTTGGCTGCTCATCATCCATCGGGCTGGAGCTTCTCTCATGATGCTG 335
QY 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
DB 336 GTGTTGGAGAAATGACAGATATCTTTGCCAAATGCAGGAAATTTAGAGATCTGATGTC 395
QY 32 GlyCysLeuValGlnThr-----AsnThrTySerPhePhe----- 43
DB 396 AACATCACTAATAGAGATGATATCAATGATACAGAGGTTCCTCATGAAATCTGGAGGAC 455
QY 44 -----ArgLeuThrLeuTyTyValGlyIleGlyValAlaAlaLeuIlePheGlyTy 61
DB 456 ATGACCAAGGATGCTTATTTATACAGTGGAAATGGTGGGTGGTGGTGGTGGTGGTGGT 515
QY 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
DB 516 ATTCAGGTTTCATTTGGTGGCTGGCAGCTGGAGGACAAATACACAAATTAGAAACAG 575
QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
DB 576 TTTTTCATGCTATAATCGCACAGGAGATAGGCTGGTTGATGTCACAGATGTTGGGAG 635
QY 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
DB 636 CTTAAACCCCGATACAGATGATGCTCTTAAGATTAATGAAGTTATTTGGTGACAAAT 695
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140

DB 696 GGAATGTTCTTTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGATTACACGT 755
QY 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
DB 756 GGTTCGAAGCTAACCTTGTGATTTGGCCATCAGTCCCTGTTCTTGGAGCTGTGACGTCT 815
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTySerLysAla 180
DB 816 GTCTGGGCAAGATATCTTCTTCTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 875
QY 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
DB 876 GGAGCAGTAGCTGAAGAGGCTTTGGCAGCAATTAGAACTGTGATTCATTTGGAGACAA 935
QY 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyAlaTrpPheTyPhe 220
DB 936 AGAAGAAGACTTGAA----- 950
QY 221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyThrGlnAsnLeuLysAsp 240
DB 951 -----AGGTACACACAAAATTTAGAAGAA 974
QY 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTy 260
DB 975 GCTAAAGAAATTTGGGATAAAGAAAGCTATTACAGCCAAATTTCTATAGGTGCTGCTTC 1034
QY 261 PhePheMetAsnGlyTyThrTyGlyLeuAlaPheTyTyThrSerLeuIleLeuAsn 280
DB 1035 CTGCTGATCATGATCTATGCTCTGGCTCTCTGGTATGGACACCTTGGTCTCTCA 1094
QY 281 GlyGluProGlyTyThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
DB 1095 GGGGAA-----TATTCATTGGCAAGTACTCACTGTATTCTTTCTGTATTAAATGGG 1148
QY 301 SerTyrcysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 120
DB 1149 GCTTTTAGTGTGGACAGCATCTCCAAGCATTTGAAGCATTTGCAAAATCAAGAGAGCA 1208
QY 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
DB 1209 GCTTATGAAATCTTCAAGATAATTGATAATTAAGCCAAATTTGACAGCTATTGCAAGAGT 1268
QY 341 GlyTyrcysProGluSerIleGlyThrValGluPheLysAsnValSerPheAsnTy 360
DB 1269 GGGCACAACACCAAGATAATTAAGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 1328
QY 361 ProSerArgProSerIleLysLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
DB 1329 CCATCTCGAAAAAGATTAAAGATCTTGAAGGGCTGAACTGAAAGGTGCAGAGTGGGCAG 1388
QY 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValGlnLeuLeuGln 400
DB 1389 ACGTGGCCCTGTTGGAAACAGTGGCTGTGGAAAGACACAAACAGTCCAGCTGATGTCAG 1448
QY 401 ArgLeuTyrcysProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
DB 1449 AGGCTCTATGACCCACACAGAGGGATGGTCAGTGTGTGATGGACAGGATATTAGACCAT 1508
QY 421 AsnValArgHisTyrcysArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
DB 1509 AATGTAAGGTTTCTACGGGAAATCATTTGGTGTGGTGTGGTGTGGTGTGGTGTGGT 1568
QY 441 ThrThrIleSerAsnAsnIleLysTyrcysArgAspAspValThrAspGluGluMetGlu 460
DB 1569 ACCACATGATGCTGAAACCAATCGTATGCGCGTGAATAATGTCCACCATGATGATGATG 1628
QY 461 ArgAlaAlaArgGluAlaAsnAlaTyrcysPheIleMetGluPheProAsnLysPheAsn 480
DB 1629 AAAGCTCTCAAGAGCCCAATGCTATGATGCTTATCATGAAATGCTCTATAAATTTGAC 1688
QY 481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500

Db	1689	ACCTGGTTGGAGAGAGAGGGCCCACTGGTTGAGTGTGGGCAAGACAGAGGAGTCCCAATT	1749
QY	501	AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu	520
Db	1749	GCACGTGCCCTGGTTTCGCAACCCCAAGATCTCTCTCTGGATGAGGCCACGTCAGCCTTG	1808
QY	521	AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThr	540
Db	1809	GACACAGAAGCGACACGCTGGTTTCAGGTGGCTCTCGGATAGAGCCAGAAAGGTCGAGACC	1868
QY	541	ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu	560
Db	1869	ACCATGTGATAGTCATCGTTGTCTACAGTTCGTAAATGCTGACGTCATCTCGTGGTTTC	1928
QY	561	LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu	580
Db	1929	GATGATGGAGTCATTCTGGAGAAGGNAATCATGATGAATCATCATGAAGAGAAAGCAATT	1988
QY	581	TyrTyrSerLeuValMetSerGln-----AspIleLysLysAlaAsp	594
Db	1989	TACTTCAAACTTGTCAACAATGCAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCT	2048
QY	595	GluGlnMetCysSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis	614
Db	2049	GATGAATCCAAAAGTGAATTCATCGCTGGAAATGTCTTCAAAATGATCAAGATCCAGT	2108
QY	615	SerVal-----LysSerIleLys---SerAppPheIleAspLysAla	627
Db	2109	CTAATTAAGAAAAGATCAACTCGTAGGAGTGTCCGTGTGATCACAAGCCCAAGACAGAAG	2168
QY	628	GluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLysIleLeu	647
Db	2169	CTTAGTACCAAGAGGCTCTCGGATGAAGTATACCTCCAGTTCTCTTTGGAGGATTATG	2228
QY	648	LysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsn	667
Db	2229	AAGCTAAATTTAACTGAATGGGCTTATTATTGTGTGTGTATTGTGGCCATTATAAAT	2288
QY	668	GlyThrValHisProValPheSerIlePheAlaLysIleIleThrMetPheGlyAsn	687
Db	2289	GGAGGCTCGCACACACNTTTCGAATATATTTTCAAAGATTATAGGGTTTTTACAAGA	2348
QY	688	---AsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIle	706
Db	2349	ATTGATGATCTCGAAACAAACACAGACAGATAGTAACTGTTTTCATTCTTCTTCTAGCC	2408
QY	707	LeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGly	726
Db	2409	CTTGGAAATTTATTTCTTTTATATCATATTTTCCCTCAGGGTTTTCACATTTGGCAAAAGCTGGA	2468
QY	727	GluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIle	746
Db	2469	GAGATCTCTACCAAGGGCTCGCATACATGGTTTTCGATCCATGCTCAGACAGAGTGTG	2528
QY	747	AlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAsp	766
Db	2529	AGTTGGTTTGATGACCTTAAACACACCACCTGAGGCATTGACTACCAAGCTCGCCAAATGAT	2588
QY	767	IleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThr	786
Db	2589	GCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTAATTAACCCAGAAATAGCA	2648
QY	787	AsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyrGluMetThrPheLeuIle	806
Db	2649	AATCTTTGGGACAGAAATAATATCTTCATCTATGTTGGCACTAACCACTGTTACT	2708
QY	807	LeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGly	826
Db	2709	TTAGCAATTGTCACCATCATTCGATAGCAGAGGTGTGTAATGAAAAATGTTGCTGGA	2768
QY	827	PheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeu	846
Db	2769	CAAGCACTGAAGATTAAGAAAGAACTAGAAAGGTCTGGAGAGATCGTACTCTGAAGCAATA	2828

Qy	847	Glus	en	lle	arg	thr	le	val	ser	leu	thr	arg	glu	lys	ala	phe	glu	glu	met	tyr	glu	866
Db	2829	GA	AA	AC	CT	CC	GA	AC	CG	CT	GT	TT	CT	TT	GA	CT	C	AG	G	A	GA	2888
Qy	867	Glu	met	leu	glu	thr	gln	his	arg	asn	thr	ser	lys	lys	ala	gln	ile	ile	gln	ser	cys	886
Db	2889	C	A	G	A	T	T	G	C	A	G	A	A	C	T	T	T	G	A	A	A	2948
Qy	887	Tyr	ala	phe	ser	his	ala	phe	le	tyr	phe	ala	tyr	ala	ala	g	ly	phe	arg	phe	g	906
Db	2949	T	T	T	C	C	T	T	C	A	C	C	C	G	A	A	T	G	A	T	T	3008
Qy	907	Tyr	leu	ile	gln	ala	g	ly	arg	met	thr	pro	gln	g	ly	met	phe	ile	val	phe	thr	926
Db	3009	T	A	C	T	T	G	T	G	T	G	C	A	T	A	A	C	T	T	G	A	3068
Qy	927	Ala	tyr	gly	ala	met	ala	ile	g	ly	lys	thr	leu	val	leu	ala	pro	gln	tyr	ser	lys	946
Db	3069	G	T	C	T	T	G	T	G	C	C	A	A	G	T	C	A	T	T	G	C	3128
Qy	947	Lys	ser	gly	ala	ala	his	leu	phe	ala	leu	leu	glu	lys	lys	pro	asn	ile	asn	ser	arg	966
Db	3129	A	A	A	T	A	T	A	T	A	T	A	T	A	T	A	T	A	T	A	T	3188
Qy	967	Ser	gln	gln	gln	lys	lys	pro	asp	thr	cys	glu	gln	g	ly	asn	leu	glu	val	ser	986	
Db	3189	A	G	C	A	G	A	G	C	C	T	A	T	G	G	A	A	G	A	T	G	3248
Qy	987	Phe	phe	tyr	pro	cys	arg	pro	asp	val	phe	ile	leu	arg	g	ly	leu	ser	leu	ser	ile	1006
Db	3249	T	T	C	A	A	C	T	A	T	C	C	A	C	C	G	A	T	C	C	A	3308
Qy	1007	Arg	g	ly	lys	thr	val	ala	phe	val	g	ly	ser	ser	g	ly	cys	g	ly	lys	ser	1026
Db	3309	A	A	G	G	C	A	G	A	C	G	T	C	T	G	G	C	A	G	A	C	3368
Qy	1027	Leu	leu	gln	arg	leu	tyr	asp	pro	val	gln	gln	gln	val	leu	phe	asn	g	ly	val	asp	1046
Db	3369	C	T	C	T	G	A	G	C	G	G	T	T	C	A	G	A	A	G	A	A	3428
Qy	1047	Lys	glu	leu	asn	val	gln	thr	leu	arg	ser	gln	ile	ala	ile	val	pro	gln	gln	pro	val	1066
Db	3429	A	A	G	C	A	T	G	A	T	G	T	C	C	A	G	A	C	A	C	T	3488
Qy	1067	Leu	phe	asn	cys	ser	ile	ala	glu	asn	ile	ala	tyr	g	ly	asp	asn	ser	arg	val	pro	1086
Db	3489	C	T	G	T	T	C	A	C	T	G	C	A	T	T	G	C	A	T	G	C	3548
Qy	1087	Leu	asp	glu	ile	lys	glu	ala	ala	asn	ala	ala	asn	ile	his	ser	phe	ile	gln	g	ly	1106
Db	3549	C	A	G	A	G	A	G	A	T	C	G	T	T	G	T	A	C	A	T	G	3608
Qy	1107	Pro	gln	lys	tyr	asn	thr	gln	val	g	ly	lys	g	ly	ala	gln	leu	ser	g	ly	c	1126
Db	3609	C	C	T	A	A	T	A	A	T	A	T	A	T	A	T	A	T	A	T	A	3668
Qy	1127	Gln	arg	leu	ala	ile	ala	arg	ala	leu	leu	gln	lys	pro	lys	ile	leu	leu	leu	asp	gln	1146

QY 1207 ArgAsnArgAspIleTyrPhelYsLeuValAsnAlaGlnSer 1220
 Db 3909 GCACAGAAAGGCATCTATTTTCAATGGTCAGTGTCCAGGCT 3950

RESULT 14

AAZ94738
 ID AAZ94738 standard; cDNA; 4646 BP.

AC AAZ94738;

DT 01-AUG-2000 (first entry)

DE Human ATP binding cassette ABCB1 (MDR1) cDNA.

KW ABCB1; ATP binding cassette; human; cholesterol; lipid disorder;
 KW atherosclerosis; lipid disorder; dyslipidemia; psoriasis;
 KW lupus erythematosus; diagnosis; gene therapy; MDR1;
 KW multidrug resistance; chromosome 7q21; ss.

XX Homo sapiens.

XX WO200018912-A2.

XX 06-APR-2000.

PF 21-SEP-1999; 99WO-EP06991.

PR 25-SEP-1998; 98US-0101706.

PA (FARB) BAYER AG.

XX Schmitz G, Klucken J;

DR WPI; 2000-293151/25.

XX Adenosine triphosphate binding proteins useful for identifying agents
 PT for treating atherosclerosis and other inflammatory disorders -

PS Claim 9; Page 110-112; 154pp; English.

XX The present sequence is that of human ATP binding cassette
 CC subfamily B protein ABCB1 cDNA. The cDNA was identified using a
 CC differential display method in which monocytes from peripheral
 CC blood were subjected to macrophage differentiation and cholesterol
 CC loading with acetylated low density lipoproteins and subsequent
 CC cleaving with high density lipoprotein (HDL3) to identify
 CC cholesterol sensitive genes. The gene maps to chromosome 7q21
 CC and is also termed MDR1 (multidrug resistance). The invention
 CC provides cholesterol-sensitive ABC genes (see AAZ94734-63). These
 CC genes, and polypeptides encoded by them, can be used for diagnostic
 CC and therapeutic applications, and for biochemical or cell-based
 CC assays to screen for pharmacologically active modulator compounds
 CC useful for the treatment of lipid disorders, atherosclerosis or
 CC other inflammatory diseases such as psoriasis and lupus
 CC erythematosus.

XX SQ Sequence 4646 BP; 1371 A; 892 C; 1129 G; 1254 T; 0 other;

Alignment Scores:

Pred. No.: 0 Length: 4646
 Score: 3521.50 Matches: 681
 Percent Similarity: 73.92% Conservative: 246
 Best Local Similarity: 54.31% Mismatches: 264
 Query Match: 57.18% Indels: 63
 DB: 21 Gaps: 10

US-09-873-409-5 (1-1222) x AAZ94738 (1-4646)

QY 1 MetIleuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
 Db 575 ATCGTGGTGGGAACCTTGGCTGCATCATCCATGGGGCTGGACCTCTCTCATGATGCTG 634

QY 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31

Db 635 GGTGTTGGAGAAATGACAGATATCTTTGCAATGACGAGAAATTTAGAGATCTGATGTC 694
 QY 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
 Db 695 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTTCATGAATCGGAGGAGAC 754
 QY 44 -----ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
 Db 755 ATGACCAGGTATGCTTATTTATACAGTGAATTTGGTGGGGTCTGCTGCTTAC 814
 QY 62 IleGlnIleSerLeuTyrIleIleThrAlaAlaAaGlnThrLysArgIleArgLysGln 81
 Db 815 ATTCAGGTTTCATTTTGGTCCCTGGCAGCTGGAAGACAAATACACAAATTTAGAAACAG 874
 QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101
 Db 875 TTTTTCATGCTATAATGCGACAGGAGATAGCGTGGTTGATGTGCACCATGTGGGGAG 934
 QY 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
 Db 935 CTTAACACCCGACTTACAGATGATGCTCTAAGATTAATGAAGTTATTGGTGACAAATTT 994
 QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
 Db 995 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGGATTACACGT 1054
 QY 141 GlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
 Db 1055 GGTGGAGCTAACCTTGTGATTTTGGCCATCAGTCCCTGTTCTGGACTGTCAGCTGCT 1114
 QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
 Db 1115 GTCTGGCAAGATACTATCTTCATTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 1174
 QY 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
 Db 1175 GGAGCAGTAGCTGAAGAGGCTCTTGGCAGCAATTAGAACTGTGATTCATTTGGAGACAA 1234
 QY 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
 Db 1235 AGAAAGAACTTGAA----- 1249
 QY 221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
 Db 1250 -----AGGTACAACACAAATAATTAGAGAA 1273
 QY 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
 Db 1274 GCTAAAAGAAATTTGGGATAAAGAAAGCTATTACAGCCAATATTCTATAGGTGCTGCTTC 1333
 QY 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsn 280
 Db 1334 CTGCTCATGTATGCATCTTATGCTCTGGCTTCTGGTATGGGACCACTTGGTCTCTCA 1393
 QY 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
 Db 1394 GGGAA-----TATTCTATTGGCAAGTACTCAGTGTATTCTTTCTGTATTAAATGGG 1447
 QY 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
 Db 1448 GCTTTTAGTGTGGACAGCATCTCCAAGCATTTGAAGCATTTTCAAAATCGAGAGAGCA 1507
 QY 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
 Db 1508 GCTTATGAATCTTCAAGATAATTGATAATAGCCCAAGTATTGACAGCTATTCCAAAGAGT 1567
 QY 341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
 Db 1568 GGGCACAACACAGATAATTAAGGAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 1627
 QY 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380

Db 1628 CCATCTCGAAAAAGAGTTAAGATCTTTGAAGGGCCCTGAAACCTGAAAGGTGCAGAGTGGGCAG 1687
QY 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
Db 1688 ACGTGGCCCTCGTTTGAACACAGTGGCTGTGGAGAGCAACACAGTCCAGCTGTATGCAG 1747
QY 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
Db 1748 AGGCTCTATGACCCACAGAGGGGATGGTCAGTGTTCATGACGAGGATATTAGGACCATA 1807
QY 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
Db 1808 AATGTAAGGTTTCTACGGGAATCATTTGGTGTGTGAGTCAGGAACCTGTATTGTTGCC 1867
QY 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
Db 1868 ACCACGATAGCTGAAACATTCGCTATGGCCGTGAAATGTCAACATGGATGAGATTGAG 1927
QY 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
Db 1928 AAGCTCTCAAGGAGGCAATGCCTATGACTTTATCATGAACCTGCCTCATAAATTGAC 1987
QY 481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
Db 1988 ACCCTGGTTGGAGAGAGGGGCCAGTTGAGTGGTGGCAGAGCAGAGGATCGCCATT 2047
QY 501 AlaArgAlaLeuValArgAsnProLysIleIleuIleLeuAspGluAlaThrSerAlaLeu 520
Db 2048 GCAGTGCCTCGTTCGCAACCCCAAGATCTCTGCTGGATGAGGCCAGCTAGCCCTTG 2107
QY 521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThr 540
Db 2108 GACACAGAGAGGAGCAGTCGTTCAAGTGGCTCTGGTAAGGCCACAAAGGTCGACC 2167
QY 541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560
Db 2168 ACCATTGTGATGCTCATCGTTTGTCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTC 2227
QY 561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580
Db 2228 GATGATGGAGTCATGTGGAGAGAGGAATCATGATGAATCATGAAGAGAGAGGCAATT 2287
QY 581 TyrTyrSerLeuLeuMetSerGln-----AspIleLysLysAlaAsp 594
Db 2288 TACTTCAAACTGTTCACAAATCCACACAGCAGGAATGAAGTTGAATTAGAAATGCAGCT 2347
QY 595 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis 614
Db 2348 GATGAATCCAAAAGTGAAATGTGATGCTTGGAAATGTCTTCAAAATGATTCAGATCCAGT 2407
QY 615 SerVal-----LysSerIleLys---SerAspPheIleAspLysAla 627
Db 2408 CTAATAAGAAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAG 2467
QY 628 GluGluSerThrGlnSerLysLeuIleSerLeuProGluValSerLeuLeuLysIleLeu 647
Db 2468 CTTAGTACCAAGAGGCTCTCGATGAAGATATACCTCCAGTTTCTTTGGAGGATTATG 2527
QY 648 LysLeuAsnLysProGluTyrProPheValValLeuGlyThrLeuAlaSerValLeuAsn 667
Db 2528 AAGCTAAATTTAACTGAATGCCCTATTATTTGTGTGTGTATTTGTGCCATTATAAAT 2587
QY 668 GlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn 687
Db 2588 GGAGGCTGCAACACGACATTTGCAATAATATTTCAAGAATATAGGGGTTTTTACAGA 2647
QY 688 ---AsnAspLysThrLeuLeuHisAspAlaGluIleTyrSerMetIlePheValIle 706
Db 2648 ATTGATGATCTGAAACAAACGACAGATAGTAATGTGTTTTCATTATTTCTAGCC 2707
QY 707 LeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGly 726
Db 2708 CTTGGAATTTATTTCTTTTATTTACATTTTCTTTCAGGGGTTTTTACATTTTGGCAAGCTGGA 2767

QY 727 GluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIle 746
Db 2768 GAGATCCTCACCAGGCGCTCCGATACATGTTTCCGATCCATGCTCAGACAGGATGTG 2827
QY 747 AlaTyrPheAspGluLysGluAsnSerThrGlyLysLeuThrThrIleLeuAlaIleAsp 766
Db 2828 AGTTGTTTGTATGACCTTAAACACCACTGGAGCATTTGACTACCAAGGCTCGCCAATGAT 2887
QY 767 IleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThr 786
Db 2888 GCTGCTCAAGTTAAAGGGCTATAGGTTTCAGGCTTCTGCTAATTTACCCAGATATAGCA 2947
QY 787 AsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyrGluMetThrPheLeuIle 806
Db 2948 AATCTTGGGACAGGAATAATTATATCTTCATCTATGTTGGCACTAAACACTGTTACTC 3007
QY 807 LeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGly 826
Db 3008 TTAGCAATTGTACCCCATCATTTGCAATAGCAGGAGTTGTTGAAATGAAATGTTGCTGGA 3067
QY 827 PheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeu 846
Db 3068 CAAGCACTGAAGATAGAAAGAACTAGAAAGTGTGGGAAGATCGCTACTGAAGCAATA 3127
QY 847 GluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGlu 866
Db 3128 GAAACTTCGGAACCGTTGTTCTTGTGACTCAGGAGCAGAAGTTTGAACATATGATGCT 3187
QY 867 GluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCys 886
Db 3188 CAGAGTTTGCAGGTACCATACAGAAAATCTCTTGGAGAAAGCACACATCTTTGGAATTACA 3247
QY 887 TyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyVala 906
Db 3248 TTTTCTTCAACCCAGGAATGATGATTTTCTATGCTGGATGTTTCCGGTTTGGAGCC 3307
QY 907 TyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIle 926
Db 3308 TACTTGGTGCCACATAAACTCATGAGCTTTGAGGATGTTCTGTAGTATTTCAGCTGTT 3367
QY 927 AlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAla 946
Db 3368 GTCTTGGTGGCCATGGCGGTGGGCAAGTCAGTTTCATTTCTCTGCTACTATGCAAGACC 3427
QY 947 LysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArg 966
Db 3428 AAAATATCAGCAGCCACATCATCATGATCATTTGAAAAAACCCTTTGATTGACAGCTAC 3487
QY 967 SerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSer 986
Db 3488 AGCAGGAAGCCCTAATCCGGAACACATTCGAAGGAATGTCACATTTGGTGAAGTTGTA 3547
QY 987 PhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGlu 1006
Db 3548 TTAACATATCCACCCGACCGGACATCCAGTGTCTCAGGACTGAGCCTGGAGGTGAAG 3607
QY 1007 ArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGln 1026
Db 3608 AAGGGCCAGACGCTGGCTCTGGTGGGCGAGGTGGCTGTGGGAAGACACAGTGTCCAG 3667
QY 1027 LeuLeuGlnArgLeuTyrAspProValGlnGlnValLeuPheAspGlyValAla 1046
Db 3668 CTCTGGAGCGGTTCTACAGACCCCTTGGCAGGGAAGTGTGCTTGTATGCAAAAGAAATA 3727
QY 1047 LysGluLeuAsnValGlnThrLeuArgSerGlnIleAlaIleValProGlnGluProVal 1066
Db 3728 AAGCGACTGAATGTTCACTGGCTCCGAGCACACCTCGTGGCATCGTGTCCAGGAGCCCAT 3787
QY 1067 LeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValPro 1086
Db 3788 CTGTTTGACTGCAGCATTTGCTGAAACATTTGCCCTATGAGACAAACACCCGGGTGTGTCA 3847


```
QY 1087 LeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeu 1106
Db 3848 CAGGAAGAGATCGTAGGGCAGCAAAAGGAGGCCAACATACATCGCTTCATCGAGTCACTG 3907
QY 1107 ProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLys 1126
Db 3908 CCTAATAATATAGCACTAAAGTAGGAGACAAGGAACCTCAGCTCTCTGGTGCCAGAAA 3967
QY 1127 GlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGlu 1146
Db 3968 CAACGCATTGCCATAGCTGTCGCCCTGTGTAGACAGCCTCATATTTGCTTTGGATGAA 4027
QY 1147 AlaThrSerAlaLeuAspAsnSerGluLysValGlnHisAlaLeuAspLysAla 1166
Db 4028 GCCACGCTGAGCTGTGATACAGAAAGTGAAGGTTGTCCAAAGACCCCTGGCAAGGCC 4087
QY 1167 ArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAsp 1186
Db 4088 AGAAGAGCGCCACCTGCATTGTGATTGCTCACCGCTGTCCACCATCCAGAAATGCAGAC 4147
QY 1187 LeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeu 1206
Db 4148 TTAATAGTGGTTTTCAGAAATGCGCAGAGTCAGAGCAGCATGGCAGCATCAGCAGCTGCTG 4207
QY 1207 ArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
Db 4208 GCACAGAAAGGCATCTATTTTCAATGGTCAGTGTCCAGGCT 4249

RESULT 15
AAD38994
ID AAD38994 standard; DNA; 4646 BP.
XX
AC AAD38994;
DT 23-SEP-2002 (first entry)
DE
XX
KW Human mdrl gene.
KW Human; hematologic malignancy; multidrug resistance; MDR; SUMO-1;
KW hypoxia inducible factor-1; small ubiquitin-like-modifier; HIF-1;
KW lymphoid disorder; chronic lymphoproliferative disorder; lymphoma;
KW myeloid disorder; lymphocytic leukaemia; thrombocythaemia; myeloma;
KW angogenic myeloid metaplasia; myeloid leukaemia; gene therapy;
KW polycythaemia vera; hypoxia responsive element; HRE; gene; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 425..4267
FT FT /*tag= a
FT FT /product= "Human MDR protein"
XX
XX WO200234291-A2.
XX
XX 02-MAY-2002..
XX
XX 25-OCT-2001; 2001WO-US49856.
XX
XX 26-OCT-2000; 2000US-243542P.
XX
XX (BGHM ) BRIGHAM & WOMENS HOSPITAL INC.
XX
XX Colgan SP;
XX
XX WPI; 2002-471427/50.
XX
XX P-PSDB; AAE24211.
XX
XX Treating a subject (at risk of) having a hematologic malignancy or
XX multidrug resistance, e.g. lymphoma or myeloma, by administering
XX hypoxia inducible factor 1 binding molecules or small
XX ubiquitin-like-modifier-1 binding molecules -
XX
XX Disclosure; Page 53-55; 92pp; English.
PS
```

```
XX The invention relates to a method of treating a subject having or at
CC risk of developing a haematologic malignancy or multidrug resistance
CC (MDR). The method involves administering hypoxia inducible factor-1
CC (HIF-1) binding molecules or small ubiquitin-like-modifier (SUMO)-1
CC binding molecules or HIF-1-SUMO-1 complex modulators. mdrl-hypoxia
CC responsive element (HRE) binding molecules or antisense nucleic
CC acid molecules and SUMO-1 binding molecules or antisense molecules
CC are useful for treating a subject having or at risk of developing
CC haematologic malignancy or MDR (e.g. a lymphoid disorder or a myeloid
CC disorder). The lymphoid disorders include lymphocytic leukaemia or
CC chronic lymphoproliferative disorders e.g. lymphoma, myeloma or chronic
CC lymphoid leukaemia. The myeloid disorders include chronic or acute
CC myeloid leukaemia, e.g. angioleukemia myeloid metaplasia, essential
CC thrombocythaemia or polycythaemia vera. The invention is used in gene
CC therapy. The present sequence is human mdrl gene.
XX
SQ Sequence 4646 BP; 1371 A; 892 C; 1129 G; 1254 T; 0 other;
```

Alignment Scores:

Pred. No.:	0	Length:	4646
Score:	3521.50	Matches:	681
Percent Similarity:	73.92%	Conservative:	246
Best Local Similarity:	54.31%	Mismatches:	264
Query Match:	57.18%	Indels:	63
DB:	24	Gaps:	10

US-09-873-409-5 (1-1222) x AAD38994 (1-4646)

```
QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 575 ATGTGGTGGGAACCTTTGGCTGCCATCATCCATGGGGCTGGAGCTCTCTCATGATGCTG 634
QY 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
Db 635 GTGTTGGAGAAATGACAGATATCTTGCAAATGCAGGAAATTTAGACAGATCTGATGTCA 694
QY 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
Db 695 AACATCACTAATAGAAGTATCATATCATGATATACAGGGTCTTTCATGAATCTGGAGGAAGAC 754
QY 44 -----ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db 755 ATGACACCGATATGCCCTATTATTATACAGTGGAAATTTGGTGGGGTGTGCTGCTGCTTAC 814
QY 62 IleGlnIleSerLeuTyrIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db 815 ATTCAGGTTTCATTTTGGTGCTGGCAGCTGGAAGCAAAATACACAAAATTAGAAAACAG 874
QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101
Db 875 TTTTTCATGCTAATAATGCAGCAGGAGATAGGCTGGTTTGTATGTGCACGATGTTGGGGAG 934
QY 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db 935 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATGAAGATTATTTGGTGCACAAAAT 994
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db 995 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGGATTTACACGT 1054
QY 141 GlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db 1055 GGTGGAAGCTAACCTTGTGATTGTTGGCCATCAGTCTCTTCTTGGAGCTGTCAGCTGCT 1114
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db 1115 GTCTGGCAAGATACATCTCTCAATTTACTGATAAAGAACTCTTAGCTATGTCAGAAAGCT 1174
QY 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db 1175 GGAGCAGTAGCTGAAGAGGCTCTTGGCAGCAATTAGAACTGTGATTGCTTGGAGGACAA 1234
```


Db 3308 TACTTGGTGGCACATAAATCAATGAGCTTTGAGGATGTTCTGTAGTATTTTCAGCTGTT 3367
QY 927 AlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAla 946
Db 3368 GTCTTTGGTGCCATGGCCCTGGGCAAGTCAGTTTCATTGCTCCTGACTATGCCAAAGCC 3427
QY 947 LysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArg 966
Db 3428 AAAATATCAGCAGCCACATCATCATGATCATTTGAAAAAACCCCTTTGATTGACAGCTAC 3487
QY 967 SerGlnGluGlyLysProAspThrCysGluGlyAsnLeuLysPheArgGluValSer 986
Db 3488 AGACGGAAGGCTTAATGCCGAACATTTGGAAGGAAATGTACATTTGGTGAAGTTGTA 3547
QY 987 PhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGlu 1006
Db 3548 TTCAACTATCCACCCAGCCGACATCCCGAGTCTTCAGGGACTGAGCCTGGAGGTGAAG 3607
QY 1007 ArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGln 1026
Db 3608 AAGGGCCAGACGCTGCTGTGTGGCAGCAGTGGCTGTGGGAAGACACAGTGGTCCAG 3667
QY 1027 LeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAla 1046
Db 3668 CTCCTGGAGCGGTTCTACGACCCCTTGGCAGGGAAGTGTCTGTATGGCAAGAAATA 3727
QY 1047 LysGluLeuAsnValGlnTyrLeuArgSerGlnIleAlaIleValProGlnGluProVal 1066
Db 3728 AAGCGACTGAATGTTTCAGTGGCTCCAGACACACCTGGGCATCGTGTCCAGGAGCCCATC 3787
QY 1067 LeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValPro 1086
Db 3788 CTGTTTACTGTCAGCATTTGCTGAGAACATTTGCCATATGGAGACAACAGCCGGGTGGTGTCA 3847
QY 1087 LeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeu 1106
Db 3848 CAGGAAGAGATCGTAGGCGCAGCAAGAGGAGCCCAACATACATGCCCTTCATCGAGTCAC 3907
QY 1107 ProGluLysTyrAsnThrGlnValGlyLeuLysGlyValAlaGlnLeuSerGlyGlyGlnLys 1126
Db 3908 CCTAATAATATAGCACTAAGTAGGAGACAAGGAACCTCAGCTCTCTGTGGCCAGAAA 3967
QY 1127 GlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuAspGlu 1146
Db 3968 CAACGATTGCCATAGCTCGTGCCCTTTGTAGACAGCCTCATATTTTGTGATGAA 4027
QY 1147 AlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAla 1166
Db 4028 GCCACGTCAGCTCTGGATACAGAAAGTGAAGGTTGTCCAAAGAGCCCTGGCAAAAGCC 4087
QY 1167 ArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAsp 1186
Db 4088 AGAAGAGCCGCACTGCTGATTGTTGCTCACCGCTGTCCACCATCCAGAAATGCAGAC 4147
QY 1187 LeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeu 1206
Db 4148 TTAATAGTGGTGTTCAGAAATGCGAGAGTCAAGAGCATGGCAGCATCAGCAGCTGCTG 4207
QY 1207 ArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
Db 4208 GCACAGAAAGGCATCTATTTTCAATGGTCAGTGTCCAGGCT 4249

Search completed: March 30, 2003, 03:16:38
Job time : 1426.3 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 30, 2003, 01:24:25 ; Search time 16520.1 Seconds
(without alignments)
2152.747 Million cell updates/sec

Title: US-09-873-409-5
Perfect score: 6159
Sequence: 1 MILGILASLVNGACPLMPL.....QELLENRDYFKLVNAQSVQ 1222

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -MODEL=frame+ p2n.model -DEV=xlh

-Q/cgn2_1/USFTO.spool/US09873409/runat_27032003.115418.19228/app_query.fasta_1.7544
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09873409 @CGN 1 1 54887 @runat_27032003.115418.19228 -NCPU=6 -ICPU=3
-NO_XLIFY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DBLOP=6 -DELEXT=7

Database :

GenEmbl:*
1: gb_ba:*
2: gb_hgt:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

RESULT 1

ALIGNMENTS

29: em_vi.*	em_htg_hum.*	31: em_htg_inv.*	32: em_htg_other.*	33: em_htg_mus.*	34: em_htg_pln.*	35: em_htg_rod.*	36: em_htg_mam.*	37: em_htg_vit.*	38: em_sy.*	39: em_htgo_hum.*	40: em_htgo_mus.*	41: em_htgo_other.*
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.												
SUMMARIES												
Result No.	Score	Match %	Query Length	DB	ID	Description						
1	6157	100.0	3702	6	AX339031	AX339031 Sequence						
2	5988.5	97.2	3621	6	AX339032	AX339032 Sequence						
3	5737	93.1	3699	6	AX478104	AX478104 Sequence						
4	5328	86.5	3177	6	AX339030	AX339030 Sequence						
5	4561.5	74.1	2856	6	AX339028	AX339028 Sequence						
6	3550.5	57.6	4286	10	CRUPGPI	M60040 C.griseus P						
7	3549.5	57.6	4304	10	CRUPGPI165	M59253 Chinese ham						
8	3533.5	57.4	4279	6	AX105082	AX105082 Sequence						
9	3533	57.4	4195	6	AX108656	AX108656 Sequence						
10	3532	57.3	4927	10	AF257746	AF257746 Rattus no						
11	3530.5	57.3	3860	6	AX322787	AX322787 Sequence						
12	3530.5	57.3	4279	6	AX105078	AX105078 Sequence						
13	3530.5	57.3	4279	6	AX105080	AX105080 Sequence						
14	3530.5	57.3	4646	6	AX391099	AX391099 Sequence						
15	3528.5	57.3	4045	12	AF269224	AF269224 Synthetic						
16	3527.5	57.3	4378	6	E02326	E02326 Multidrug r						
17	3526	57.2	3858	4	OAU78609	U78609 Ovis aries						
18	3525.5	57.2	4279	6	AX105057	AX105057 Sequence						
19	3525.5	57.2	4317	4	AF045016	AF045016 Canis fam						
20	3525.5	57.2	4317	6	AX105059	AX105059 Sequence						
21	3524.5	57.2	4186	6	AX108654	AX108654 Sequence						
22	3521.5	57.2	3860	6	AX332789	AX332789 Sequence						
23	3521.5	57.2	4646	6	AX336420	AX336420 Sequence						
24	3521.5	57.2	4646	6	AX336708	AX336708 Sequence						
25	3521.5	57.2	4646	6	I49610	I49610 Sequence 2						
26	3521.5	57.2	4646	9	HUMMDR1	M14758 Homo sapien						
27	3521.5	57.2	4669	6	I08557	I08557 Sequence 3						
28	3521.5	57.2	6505	6	AR028671	AR028671 Sequence						
29	3521.5	57.2	8630	6	AX012320	AX012320 Sequence						
30	3521.5	57.2	8630	6	AX012321	AX012321 Sequence						
31	3521.5	57.2	9318	6	AR028672	AR028672 Sequence						
32	3520	57.2	4264	6	AR051647	AR051647 Sequence						
33	3520	57.2	4264	6	AR051650	AR051650 Sequence						
34	3518	57.1	4323	10	AF286167	AF286167 Rattus no						
35	3514	57.1	3840	6	AX481416	AX481416 Sequence						
36	3514	57.1	3988	6	AX024454	AX024454 Sequence						
37	3514	57.1	4192	9	AF016535	AF016535 Homo sapi						
38	3505.5	56.9	4280	10	CRUPGPII	M60041 C.griseus P						
39	3503.5	56.9	4669	6	AR091275	AR091275 Sequence						
40	3503.5	56.9	4669	6	AR203322	AR203322 Sequence						
41	3503	56.9	4189	6	AX322791	AX322791 Sequence						
42	3503	56.9	4298	10	MUSMDR	M14757 Mouse multi						
43	3499.5	56.8	3834	4	CFA419568	AJ419568 Canis fam						
44	3499	56.8	3888	10	AY082609	AY082609 Rattus no						
45	3499	56.8	4788	6	AX322793	AX322793 Sequence						

AX339031
LOCUS AX339031 3702 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 13 from Patent WO0194400.
ACCESSION AX339031
VERSION AX339031.1 GI:18129123
KEYWORDS human.
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Frank, M.H. and Sayegh, M.H.
TITLE A gene encoding a multidrug resistance human p-glycoprotein
homologue on chromosome 7p15-21 and uses thereof
JOURNAL Patent: WO 0194400-A 13 13-DEC-2001;
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES
source
1. 3702
/organism="Homo sapiens"
/db xref="taxon:9606"
BASE COUNT 1098 a 708 c 856 g 1039 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 3702
Score: 6157.00 Matches: 1222
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.97% Indels: 0
DB: 6 Gaps: 0
US-09-873-409-5 (1-1222) x AX339031 (1-3702)
Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 34 ATGATCTGGGTATACGGCATCTGGTCAATGGAGCCCTGCCTTCCTTAATGCCACTG 93
Qy 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
Db 94 GTTTTAGAGGAATGAGTGATTAACCTATTAGTGGATGTCTAGTCCAAACTACACATAC 153
Qy 41 SerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGly 60
Db 154 TCCTTCTTCAGGTTGACCCGTATATATGTTGGAATAGGTGTGCTGCTTCGATTTTGGT 213
Qy 61 TyrIleGlnIleSerLeuTyrIleThrAlaAlaArgGlnThrLysArgIleArgLys 80
Db 214 TACATACAGATTTCTCTGTGGATTATTAACATGCACACACAGACCAAGAGATTTCGAAA 273
Qy 81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGly 100
Db 274 CAGTTTTTTTCATTTCAGTTTTGGCACAGACATCGGCTGGTTGTATAGCTGTGACATCGGT 333
Qy 101 GluLeuAsnThrArgMetThrAspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db 334 GAACCTTAACACTCGCATGCACAGACATTGACAAATCAGTCAGTGATGTTGGAGATAAGATT 393
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db 394 GCTCTGTTGTTTCAAAAACATGTCCTATTTTCGATTGGCTGGCAGTTGGTTGGTGAAG 453
Qy 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db 454 GGCTGGAAACTCACCTTAGTGACTCTATCCACGTCCTCTCTTATTAATGGCTTCAGCGGCA 513
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db 514 GCATGTTCTAGGATGGTCATCTCATTCACAGTAAGGAATTAAGTGCCCTATTCCAAAGCT 573
Qy 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db 574 GGGGCTGTGGCAGAGAAGTCTTGTCATCATCCGAAACAGTCATAGCCTTTTAGGGCCAC 633

Qy 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
Db 634 GAGAAAGAACTTCAAAGGCTTTTCCCTTTTAAATATAACAAGATATGCTTGGTTTATTTT 693
Qy 221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
Db 694 CCCAGTGGCTACTAAGTTGTGTCTGTNTTGTAGGTATACACAGAATCTCAAGAT 753
Qy 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
Db 754 GCAAGATTTTGGCATAAAGAGCTATAGCTTCAAAGTGCTCTTGGTGTGTGTATC 813
Qy 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
Db 814 TTCTTTATGAATGGAACCTATGACCTTGTCTTTTGGTATGGAACCTCTTGTATTCCTAAT 873
Qy 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
Db 874 GGAGAACTCGATATACCATCGGAGCTGTCTTGTCTGTCTTTAGTGTAAATCCATAGC 933
Qy 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
Db 934 AGTTATTGCANTGGAGCAGACGTCCTCACTTTGAAACCTTCGCAATAGCCCGAGAGCT 993
Qy 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
Db 994 GCCTTTTCATATTTTCCAGGTTATTGATAGAAACCCAGTATAGATAACTTTTCCACAGCT 1053
Qy 341 GlyTyrLysProGluSerIleGlyThrValGluPheLysAsnValSerPheAsnTyr 360
Db 1054 GGATATAAACCTGAATCCATAGAGGAACCTGTGGAATTTAAAAATGTTCTTCAATAT 1113
Qy 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
Db 1114 CCATCAAGACCATCTATCAGATTCTGAAGGTCCTGAATCTCAGAAATTAAGTCTGGAGAG 1173
Qy 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
Db 1174 ACAGTGCCTTGTGCTGCTCAATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAG 1233
Qy 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
Db 1234 AGGTTATATGATCCGATGATGCTTTTATCATGCTGGATGAGAATGACATCAGAGCTTTA 1293
Qy 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
Db 1294 AATGTGGCGGATTTATCGAGACCATTTGGAGTGTGTAGTCAAGAGCTGTTTGTTCGGG 1353
Qy 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
Db 1354 ACCACCATCAGTAACAATATCAAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAG 1413
Qy 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
Db 1414 AGAGCAGCAAGGAAGCAATGCGTATGATTTTATCATGAGGTTTCTCTAATAATTTAAT 1473
Qy 481 ThrLeuValGlyGlyLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
Db 1474 ACATTGTGGGAAAGAGAGCTCAATAGTGGAGGGCAGAAACAGAGATTCGCANTT 1533
Qy 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
Db 1534 GCTCGTGCTTAGTTCGAAACCCCAAGATTCTCATTTTAGATGAGGCTAGCTCTGCCCTG 1593
Qy 521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThr 540
Db 1594 GATTTCAGAAAGCAAGTCAGCTGTTTCAAGCTGCACCTGGAGAAGCGCAGCAAGGTCCGACT 1653
Qy 541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560
Db 1654 ACATCGTGTAGCAGACCCGACTTCTTACTATTTCGAAGTGCAGATTGTGATGGACCTTA 1713
Qy 561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580

Db 1714 AAGGATGGGAATGTCGGGAGAAAGAGACATGCTGAACATAATGGCAAAACAGAGTCTA 1773
Qy 581 TTTTyrSerLeuValMetSerGlnAspIleLeuLysAlaAspGluGlnMetGluSerMet 600
Db 1774 TATTAATCACTTGTGATGTCACAGGATATATAAAGAGTGATGAAACAGATGGAGTCAATG 1833
Qy 601 ThrTySerThrGluArgIysThrAsnSerLeuProLeuHisSerValIysSerIleLys 620
Db 1834 ACATAATCTTACTGAAAGAAAGACCAACTCACTTCCCTGTCACATCTGTGAGAGCATCAAG 1893
Qy 621 SerAspPheIleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGlu 640
Db 1894 TCAGACTTCAATGACAGGCTGAGGAATCCACCACATCTAAAGAGATAAGTCTTCCTGAA 1953
Qy 641 ValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGly 660
Db 1954 GTCCTCTATAAANAATTTAAGATTAAACAGCCTGAATGGCCTTTGTGGTGTGGGG 2013
Qy 661 ThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleLeuPheAlaLys 680
Db 2014 ACATTGGCTTCTGTTCTAATGGAACGTGTTCAATCCAGTATTTTCCATCATCTTTGCAAAA 2073
Qy 681 IleIleThrMetPheGlyAsnAsnAspLysThrThreuleyshiAspAlaGluIleTy 700
Db 2074 ATTATTAACCATGTTTGGNAATATATGATAAACACACATTAAGCATGATGAGAAATTTAT 2133
Qy 701 SerMetIlePheValIleLeuGlyValIleCysPheValSerTyPheMetGlnGlyLeu 720
Db 2134 TCCATGATATTCGTCATTTGGTGTTATTTGGTCTTTGTCAGTATTTTCATGCAAGGATTA 2193
Qy 721 PheTyGlyVargAlaGlyIleLeuThrMetArgLeuArgHisLeuAlaPheLysAla 740
Db 2194 TTTTACGGCAGACAGGGGAAATTTTAAACGATGAGATTAAAGACACTTGGCCTTCAAAGCC 2253
Qy 741 MetLeuTyGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThr 760
Db 2254 ATGTTATATCAGGATATGTCCTGGTTGATGAAAGGAAACAGCACAGGAGGCTTGACA 2313
Qy 761 ThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyVal 780
Db 2314 ACATAATTAGCCATAGATATAGCACAAATTAAGGAGCAACAGGTTCCAGGATGGGTC 2373
Qy 781 LeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyGlyTrp 800
Db 2374 TTAACCAAAATGCAATCAATGAGCTTTTCAGTTATCATTTCTTATATATGATGATG 2433
Qy 801 GluMetThrPheLeuLeuLeuSerIleAlaProValLeuAlaValThrGlyMetIleGlu 820
Db 2434 GAGATGACATTCCTGATTTCTGAGTATTTGCTCCAGTACTTGGCCGTGACAGGAATGATGAA 2493
Qy 821 ThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLys 840
Db 2494 ACCGACGAAATGACTGGAATTTGCCAAACAAAGATAAGCAAGACTTAAGCATGCTGGAAG 2553
Qy 841 IleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAla 860
Db 2554 ATAGCAACTGAACCTTTGGAGATATACGTACTATAGTGTCATTAAACAGGGAAGGCC 2613
Qy 861 PheGluGlnMetTyGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAla 880
Db 2614 TTCGAGCAATGATGAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCA 2673
Qy 881 GlnIleIleGlySerCysTyPheAlaPheSerHisAlaPheIleTyPheAlaTyAlaAla 900
Db 2674 CAGATTAATGGAAGCTGTTATGATTCATTCAGCCATGCGCTTTTATATATTTGCTATGAGCA 2733
Qy 901 GlyPheArgPheGlyAlaTyLeuIleGlnAlaGlyArgMetThrProGluGlyMetPhe 920
Db 2734 GGGTTTCGATTGGAGCTTATTAATTCAGCTGGAGCAATGACCCAGAGGCGCATGTTTC 2793
Qy 921 IleValPheThrAlaIleAlaTyGlyAlaMetAlaIleGlyLysThrLeuValLeuAla 940

Db 2794 ATAGTTTTTACTGCAATTCATATGAGAGCTATGGCCATCGGAAAAACGCTCGTTTGGCT 2853
Qy 941 ProGluTySerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLys 960
Db 2854 CCGTAATATTCCAAAGCCAAATCGGGGGCTGCGCATCTGTTGGCTTTGTTGAAAAAGAAA 2913
Qy 961 ProAsnIleAspSerArgSerGlnGluLysLysProAspThrCysGluGlyAsnLeu 980
Db 2914 CCAAATATATAGACGGCGAGTCAAGAAAGGAAAGAAAGCCAGACACATGTGAAGGAATTTA 2973
Qy 981 GluPheArgGluValSerPhePheTyPProCysArgProAspValPheIleLeuArgGly 1000
Db 2974 GAGTTTCGAGAAGTCTCTTTCTTATCCATGTGCGCCAGATGTTTTCATCTCCGTGGC 3033
Qy 1001 LeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGly 1020
Db 3034 TTATCCCTCAGTATTGAGCGAGGAAAGACAGTAGCATTTTGTGGGAGCAGCGCTGTGG 3093
Qy 1021 LysSerThrSerValGlnLeuLeuGlnArgLeuTyPAspProValGlnGlyGlnValLeu 1040
Db 3094 AAAAGCACTTCTGTCACTTCTGCAGAGACTTTATGACCCCGTGCAGGACAAAGTGCTG 3153
Qy 1041 PheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIle 1060
Db 3154 TTGTATGGTGTGATGCCAAAGAAATTTGAATGTACAGTGGCTCCGTTCCCAATATGCAATC 3213
Qy 1061 ValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyGlyAsp 1080
Db 3214 GTTCTCTCAAGAGCCTGTGCTTCACTGACGATTTGCTGAGAACATCGCTATGTTGAC 3273
Qy 1081 AsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaIleHis 1100
Db 3274 AACAGCCGTGTGTGCTATGATGATGATCAAGAAAGCCGCAAAATGCAGCAATATCCAT 3333
Qy 1101 SerPheIleGluGlyLeuProGluLysTyPAsnThrGlnValGlyLysGlyAlaGln 1120
Db 3334 TCTTTTATTGAAGGCTCTCCCTGAGAAATACACACAAGTTGGACTGAAAGAGGACAG 3393
Qy 1121 LeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLys 1140
Db 3394 CTTTCTGGCGGCAGAAACAAAGACTAGCTATTTGCAAGGGCTCTTCCCAAAACCCAAA 3453
Qy 1141 IleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGln 1160
Db 3454 ATTTTATTGTTGGATGAGGCCACTTCAGCCCTCGATAATGACAGTGAAGAGTGGTTTCA 3513
Qy 1161 HisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSer 1180
Db 3514 CATGCCCTTGTATAAGCCAGGACGGGAAGGACATGCTAGTGGTCACTCACAGGCTCTCT 3573
Qy 1181 AlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGly 1200
Db 3574 GCATTCAGAACGCGATTTGATGTTGTTCTGCACAATGGAAAGATAAAGGAACAGGA 3633
Qy 1201 ThrHisGlnGluLeuLeuArgAsnArgAspIleTyPheLysLeuValAsnAlaGlnSer 1220
Db 3634 ACTCATCAAGAGCTCTCGAGAAATCGACATATATTTAAAGTTAGTGAATGCACAGTCA 3693
Qy 1221 ValGln 1222
Db 3694 GTGCAG 3699

RESULT 2

AX339032

LOCUS AX339032 Sequence 14 from Patent WO0194400.

DEFINITION AX339032

ACCESSION AX339032

VERSION AX339032.1 GI:18129124

KEYWORDS human.

SOURCE Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

3621 bp DNA linear

PAT 09-JAN-2002

REFERENCE 1
AUTHORS Frank, M.H. and Savegh, M.H.
TITLE A gene encoding a multidrug resistance human p-glycoprotein
homologue on chromosome 7p15-21 and uses thereof
JOURNAL Patent: WO 0194400-A 14 13-DEC-2001;
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES Location/Qualifiers
source 1..3621
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 1081 a 696 c 842 g 1002 t
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 3621
Score: 5988.50 Matches: 1195
Percent Similarity: 97.79% Conservative: 0
Best Local Similarity: 97.79% Mismatches: 0
Query Match: 97.23% Indels: 27
DB: 6 Gaps: 1
US-09-873-409-5 (1-1222) x AX339032 (1-3621)
QY 1 MetIleuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
DB 34 ATGATCTGGGTATACCTGGCATCACTGGTCAATGGAGCGCTTCCTTTAATGCCACTG 93
QY 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
DB 94 GTTTTAGGAGAAATAGTGATTAACCTTATTAGTGGATGTCTAGTCCAAACTAACACATAC 153
QY 41 SerPhePheArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePheGly 60
DB 154 TCITTCACAGTTGACCCCTGATTATAGTTGGAAATAGGTGTGCTGCCCTTGATTTTGGT 213
QY 61 TyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLys 80
DB 214 TACATACAGATTTCCTGTGGATTATTAACATGCAGCACGACAGCAAGAGGATTCGAAAA 273
QY 81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGly 100
DB 274 CAGTTTTTTTCATTTCAGTTTTGGCACAGGACATCGGCTGGTTGATAGCTGTGACATCGGT 333
QY 101 GluLeuAsnThrArgMetThrAspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
DB 334 GAACCTTAACACTCCATGCACAGACATTCACAAATCACTGATGGTATGGAGATAAGATT 393
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
DB 394 GCTCTGTTGTTTCAAAACATGCTCTACTTTTTCGATTGGCCTGGCAGTTGTTGGTGAAG 453
QY 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProIleuIleMetAlaSerAlaAla 160
DB 454 GGCTGGAACCTCACCTAGTGACTCTATCCACGCTCCTCTTATAATGGCTTCAGCGGCA 513
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
DB 514 GCATGTTCTAGGATGGTCATCTCATTCACAGTAAGGAATTAAGTGCCTTATTCGAAAGCT 573
QY 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
DB 574 GGGGCTGTGGCAGAGAAGCTTGTGCATCAATCCGACAGCTCATAGCTTTAGGGCCACG 633
QY 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
DB 634 GAGAAAGAACTTCAA----- 648
QY 221 ProGlnTrpLeuLeuSerCysValLeu**PheValArgTyrThrGlnAsnLeuLysAsp 240
DB 649 -----AGGTATACAGAAATCTCAAGAT 672
QY 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260

DB 673 GCAAGGATTTTGGCATAAAAAAGGACTATAGCTTCAAAAGTGTCTCTTGGTGTGTGTAC 732
QY 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsn 280
DB 733 TTTCTTTATGAATGGAAACCTATGAGCTTGTCTTTTGGTATGGAAACCTCTTGTATCTTAAT 792
QY 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
DB 793 GGAGAACCTGGATATACCATCGGACTGTCTTGTCTGTCTTTCTTTAGTGTAAATCCATAGC 852
QY 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
DB 853 AGTTATTGCAATTGGAGCAGCAGTCCCTCACTTTGAAACCTTCGCAATAGCCGAGAGCT 912
QY 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
DB 913 GCCTTTTCATATTTTCAGGTTTATGATAAGAAACCCAGATATAGATNACTTTCCACAGCT 972
QY 341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
DB 973 GGATATAAACCTGAATCCATAGAGGAACCTGTGGAATTTTAAAAATGTTCTTCTTCAATTAT 1032
QY 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
DB 1033 CCATCAAGACCATCTATCAAGATTCTGAAAGGTCTGAATCTCAGAAATTAAGTCTGGAGAG 1092
QY 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
DB 1093 ACAGTCGCCCTTGTGCGTCTCAATGGCAGTGGAGAGTACGGTAGTCCAGCTTCTGCAG 1152
QY 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
DB 1153 AGGTTATATGATCCGATGATGCTTATCATGGTGGATGAGATGCATCAGAGCTTTA 1212
QY 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
DB 1213 AATGTGCGCAATTATCGAGACCATATTGGAGTGGTTAGTCAAGAGCCTGTTTGTTCGGG 1272
QY 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
DB 1273 ACCACCATCAGTAACAAATATCAAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAG 1332
QY 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
DB 1333 AGAGCAGCAAGGGAACAAATGCTATGATTTTATCATGGAGTTTCTTAATAAATTAAT 1392
QY 481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
DB 1393 ACATTGTTAGGGGAAAAAGGAGCTCAAAATGAGTGGAGGCGAGAAACAGAGGATCGCAAT 1452
QY 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
DB 1453 GCTCTGTCCTTAGTTGCAACCCCAAGATTTCTGATTTTATGATGAGGCTAGCTCTGCCCTG 1512
QY 521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThr 540
DB 1513 GATTGAAAGCAAGTCAGCTGTTCAAGCTGCATCGAGAGCGGAGCAAGGTCGACT 1572
QY 541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560
DB 1573 ACAATCGTGTAGTACACCGACTTCTTACTATTTCGAAGTGCAGATTTGATTTGACCCCTA 1632
QY 561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580
DB 1633 AAGGATGGAAATGCTGGCGAGAAAGGAGCAGCATGCTGAATTAATGGCAAAACAGGCTCTA 1692
QY 581 TyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMet 600
DB 1693 TATTATTCACTTGTGATGTCACAGGATATTAAAAAGCTGATCAACAGATGGAGTCAATG 1752
QY 601 ThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLys 620
DB 1753 ACATATTCTACTGAAAGAAAGACCAACTCACTTCTCTGCACTCTGTGAGAGCATCAAG 1812

FEATURES Location/Qualifiers

source
1. .3699
/organism="Homo sapiens"
/db_xref="taxon:9606"
/notes="Incyte ID No: 7472030CB1"

BASE COUNT 1116 a 707 c 860 g 1016 t

ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 3699
Score: 5737.00 Matches: 1156
Percent Similarity: 93.78% Conservative: 4
Best Local Similarity: 93.45% Mismatches: 9
Query Match: 93.15% Indels: 68
DB: 6 Gaps: 4

US-09-873-409-5 (1-1222) x AX478104 (1-3699)

QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
DB 145 ATGATCTGGGTACTACTGACATCACTGTTCAATGGAGCCTTCCTTCTTAATGCCACTG 204
QY 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
DB 205 TGTATAGGAAATAGTGATAACCTTATTAGTGATGTCTAGTCCACACTAACACA 264
QY 41 SerPhePheArg-----LeuThr 46
DB 265 AATTATCAGAACTGTACTAGTCTCAAGAGAGCTGAATGAAGATATGACTCTGTTGACC 324
QY 47 LeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeu 66
DB 325 CTGTATTATGTTGGAATAGGTGTTGCTGCTTGAATTTTGGTTACATACAGATTTCCCTG 384
QY 67 TrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGlnPhePheHisSerVal 86
DB 385 TGGATTATAACTGCAGCAGCAGACAGCAAGAGGATTCGAAACAGTTTTTTCATTGAGTT 444
QY 87 LeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGluLeuAsnThrArgMet 106
DB 445 TTGSCACAGGACATCGGCTGGTTTGAAGTGTGACATCGGTGAACCTTAACATCGGATG 504
QY 107 Thr---AspIleAspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGln 125
DB 505 ACAGATCAGATTCACAAATCAGTGATGGTATTGGAGATAAGATTGCTGTTGTTTCAA 564
QY 126 AsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLysGlyTyrLysLeuThr 145
DB 565 AACATGTTACTTTTTCGATTGGCTGGCAGTTGTTGTTGGTGAAGGCTCGAAACTCACC 624
QY 146 LeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAlaCysSerArgMet 165
DB 625 CTAGTGACTCTATCCAGCTCTCTCTTATATATGCTTCAGCGGAGCATGTTCTAGATG 684
QY 166 ValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGlu 185
DB 685 GTCATCTCATGACCATTAAGGATTAAGTGCCTATTCCAAAGCTGGGGCTGTGGCAGAA 744
QY 186 GluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGln 205
DB 745 GAAGTCTTGTTCATCAATCCGAAACAGTATAGCTTTAGGCCCCAGGAGAAAGAACTTCAA 804
QY 206 ArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeu 225
DB 804 ----- 804
QY 226 SerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 245
DB 805 -----AGGTATACACAGAAATCTCAAGAGTGCAGAGGATTTTGGC 843
QY 246 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGly 265
DB 844 ATAAAAAGGACTATAGCTTCAAAAAGTGTCTCTGTTGGTGTGTGTACTTCTTTATGAATGA 903

QY 266 ThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 285
DB 904 ACCTATGGACTTGGCTTTTGGTATGGAACTCTCTGATTCTTAATGGAGAACTGGATAT 963
QY 286 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 305
DB 964 ACCATCGGAGCTGTTCTTCTGTTCTTTTAGTGTAAATCCATAGTAGTATTATTGCAITGGA 1023
QY 306 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 325
DB 1024 GCAGCAGTCCCTCACCTTTGAAACCTTCGCAATAGCCGAGGAGCTGCCCTTCATATTTTC 1083
QY 326 GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 345
DB 1084 CAGTTATTGATAAGAAACCCAGTATAGGTAACTTTTCCACAGCTGGATATAAACCTGAA 1143
QY 346 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 365
DB 1144 TCCATAGAAGAACTGTGGAATTTAAAAATGTTCTTCTCAATTTATCCATCAAGGCCATCT 1203
QY 366 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 385
DB 1204 ATCAAGATTCTGAAAGGTCTGAAATCTCGGAATTAAGTCTGGAGAGACAGTCGCTTGGTC 1263
QY 386 GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro 405
DB 1264 GGTCTCAATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAGAGGTATATGATCCG 1323
QY 406 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 425
DB 1324 GATGATGGCTTTATCATCTGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGCATAT 1383
QY 426 ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn 445
DB 1384 CGAGACCATATTGGAGTGGTTAGTCAAGAGCCTGTTTGTTCGGGACCACCATCAGTAAC 1443
QY 446 AsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGlu 465
DB 1444 AATATCAAGTATGGACAGATGATGTGACTGATGAAGAGATGGAGAGACAGCAAGGAA 1503
QY 466 AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 485
DB 1504 GCAAAATCGCTATGATTTTATCATGGAGTTTCTTAATAAATTTAATACATTTGGTAGGGAA 1563
QY 486 LysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal 505
DB 1564 AAAGGAGCTCAATGATGAGTGGAGGCGAGAAACAGAGATCGCAATTGCTGCTGCTTAGTT 1623
QY 506 ArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 525
DB 1624 CGAAACCCCAAGATTTCTGATTTTAGATGAGGCTACGCTGCCCCCTGGATTCAGAAAGCAAG 1683
QY 526 SerAlaValGlnAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAla 545
DB 1684 TCAGCTGTTCAAGCTGCACCTGGAGAAAGGCGAGCAAGGTCGAGACTCAATTCGTGGTAGCA 1743
QY 546 HisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeu 565
DB 1744 CACCGACTTCTACTATTGAAAGTGCAGATTTGATTTGTGACCCCTAAAGGATGAATGCTG 1803
QY 566 AlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuVal 585
DB 1804 CGCGAGAAAGGAGCGCATGCTGAACCTAATGGCAAAACAGAGTCTATATATTCACTTGTG 1863
QY 586 MetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu 605
DB 1864 ATGTCACAGGATATTAAGCTGATGAACAGATGGAGTCAATGACATATTCTACTGAA 1923
QY 606 ArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIleAsp 625
DB 1924 AGAAAGCAACCACTCACTTCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTCATTGAC 1983


```
QY 626 LyeAlaGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLys 645
DB 1984 AAGCGTGAGGAATCCACCCCAATCTAAAGAGATAAGTCTTCCCTGAAGTCTCTCTATTAAAA 2043
QY 646 IleLeuLysLeuAsnLysProGluTppProPheValValLeuGlyThrLeuAlaSerVal 665
DB 2044 ATTTTAAAGTTAAACAAGCCTGAATGCCCTTTTGTGGTCTCGGGACATTTGGCTTCTGTT 2103
QY 666 LeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPhe 685
DB 2104 CTAAATGGAACTGTTTCATCCAGTATTTCATCATCTTTGCAAAAAATTATAACCATGTTT 2163
QY 686 GlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheVal 705
DB 2164 GGAATAATGATAAAACCAACCATTAAGAGCATGATGCAGAAATTTATTCCATGATATTTCGTC 2223
QY 706 IleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAla 725
DB 2224 ATTTTGGGTGTTATTGCTTTTGTCTAGTTATTTCATG-----CATG----- 2259
QY 726 GlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAsp 745
DB 2260 -----CAGGAT 2265
QY 746 IleAlaTppPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIle 765
DB 2266 ATTGCTGTGTTTGTATGAAAGGAAAAACAGCACAGGAGGCTTGACAAACAATATTAGCCATA 2325
QY 766 AspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAenAla 785
DB 2326 GATATAGCAAAATTCAGAGGCACAGGTTCCAGGATTTGGCGTCTTTAAACAAAAATGCA 2385
QY 786 ThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTppGluMetThrPheLeu 805
DB 2386 ACTAACATGGGACTTTTCAGTTATCATTTCTTTATATATGATGGAGATGACATTTCTCTG 2445
QY 806 IleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaIleMetThr 825
DB 2446 ATTTCTGAGTATTGCTCCAGTACTTGCCTGTCAGAGGAATGATTGAAACCCGACGATGACT 2505
QY 826 GlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAla 845
DB 2506 GGATTTGCCAACAAAGATAAAGCAAGCACTTAAGCATGCTTGGAAAGATAGCAACTGAAGCT 2565
QY 846 LeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyr 865
DB 2566 TTGGAGAAATATACGTACTATAGTGTCATTAAACAGGGAAGAGCCCTTCAGCAAAATGAT 2625
QY 866 GluGluMetLeuGlnThrGlnHisArgAsnThrSerLysAlaGlnIleIleGlySer 885
DB 2626 GAAGAGATGCTTCAGACTCAACACAGAAAATACCTCGAAGAAAGCACAGATTATTGGAAGC 2685
QY 886 CysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGly 905
DB 2686 TGTATTGCAATTGAGCCATGCCCTTTATATATTGTGCTATGCGCAGGGTTTCGATTGGA 2745
QY 906 AlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAla 925
DB 2746 GCCTATTTAATTCAGCTCGAGCAATGACCCAGAGGGCATGTTCTAGATTTTACTGCA 2805
QY 926 IleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLys 945
DB 2806 ATTGCAATATGGAGCTATGGCCATCGGAAACCGCTCGTTTGGCTCCTGAATATTCCAAA 2865
QY 946 AlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSer 965
DB 2866 GCCAAATCGGGGGCTGCGCATCTGTTTGGCTTGTGGAAAAAGAAACCAATATAGACAGC 2925
QY 966 ArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluVal 985
DB 2926 CCGAGTCAAGAGGGGAAAGCCAGACATGTGAAGGAAATTTAGAGTTTCAGAGATC 2985
QY 986 SerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIle 1005
```

```
DB 2986 TCTTTTCTTATCCATGTCGCCAGATGTTTTTCATCTCCGTGGCTTATCCCTCAGTATT 3045
QY 1006 GluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerVal 1025
DB 3046 GAGCGAGGAAAGACAGTAGCATTTTGTGGGAGCAGCGGCTGTGGGAAAGACATTTCTGTT 3105
QY 1026 GlnLeuLeuGlnArgLeuTyrAspProValGlnGlnValLeuPheAspGlyValAsp 1045
DB 3106 CCACTTCTCAGAGACTTTATGACCCCGTCAAGGACAAAGTGTGTTTGTATGTTGGAT 3165
QY 1046 AlaLysGluLeuAsnValGlnTppLeuArgSerGlnIleAlaIleValProGlnGluPro 1065
DB 3166 GCAAAAGAAATTGAATGTACAGTGGTCCGTTCCTCAATAGCAATCGTTCTCAAGAGCCT 3225
QY 1066 ValLeuPheAsnCysSerIleAlaGluAenIleAlaTyrGlyAspAsnSerArgValVal 1085
DB 3226 GTGCTCTTCACTGCAGCATTTGCTGAGACATCGCTATGTTGACAAACGCGGTGGTG 3285
QY 1086 ProLeuAspGluIleLysGluAlaAlaAenAlaAlaAenIleHisSerPheIleGluGly 1105
DB 3286 CCATTAGATGAGATCAAGAGCCGCAATGCGCAATATCCATTTCTTTATTGAAGGT 3345
QY 1106 LeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGln 1125
DB 3346 CTCCCTGAGAAATACAAACACCAAGTTGGACTGAAAGGAGCACAGCTTTCTGCGGCCAG 3405
QY 1126 LysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuAsp 1145
DB 3406 AAACAAAGACTAGCTATTGCAAGGCTCTTCTCCAAAACCCAAATTTTATTGTTGGAT 3465
QY 1146 GluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLys 1165
DB 3466 GAGGCCACTTCAGCCCTCGATAATGACAGTGAAGAGTGGTTTCAGCATGCCCTTGATAAA 3525
QY 1166 AlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAenAla 1185
DB 3526 GCCAGACGGGAAGGACATGCTTAGTGTCCTCAGGCTCTCTGCAATTTCAGAACGCA 3585
QY 1186 AspLeuIleValValLeuHisAsnGlyLysIleLysValGlnGlyThrHisGlnGluLeu 1205
DB 3586 GATTTGATAGTGGTCTGCACAATGGAAGATGAAGAAACAAGGAACCTCATCAAGAGCTC 3645
QY 1206 LeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 1222
DB 3646 CTGAGAAATCGAGACATATATTTTAAGTTAGTGAATGCACAGTCAGTGCAG 3696
RESULT 4
AX339030 3177 bp DNA linear PAT 09-JAN-2002
LOCUS Sequence 12 from Patent WO0194400.
DEFINITION AX339030
ACCESSION AX339030
VERSION AX339030.1 GI:18129122
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Frank,M.H. and Sayegh,M.H.
TITLE A gene encoding a multidrug resistance human p-glycoprotein
homologue on chromosome 7p15-21 and uses thereof
JOURNAL Patent: WO 0194400-A 12 13-DEC-2001;
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES
source
1. .3177
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 970 a 601 c 735 g 870 t 1 others
ORIGIN
Alignment Scores: 0 Length: 3177
Pred. No.:
```


Score: 5328.00 Matches: 1058
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 86.51% Indels: 0
DB: 6 Gaps: 0

US-09-873-409-5 (1-1222) x AX339030 (1-3177)

QY 165 MetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAla 184
DB 1 ATGGTCATCTCAATGACAGTAAGAAATTAAGTCCCTATTCCAAAGCTGGGGCTGGCA 60
QY 185 GluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeu 204
DB 61 GAAGAAGCTTGTCAATCAATCCGAACAGTCATAGCCCTTAGGGCCCCAGGAGAAAGAACTT 120
QY 205 GlnArgSerPheLeuLeuAenIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeu 224
DB 121 CAAGAAGTCTTTCTCTTTAAATATATAACAAGATATGCTTTGGTTTATTTTCCCCAGTGGCTA 180
QY 225 LeuSerCysValLeu***PheValArgTyrThrGlnAenLeuLysAspAlaLysAspPhe 244
DB 181 CTAAGTTGTGTCTGTTNTTGTAAAGGTATACACAGAAATCTCAAGAATGCCAAGGATTTT 240
QY 245 GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsn 264
DB 241 GGCATAAAAGGACTATAGCTTCAAAAGTGTCTCTTGGTGTGTACTTCTTTATGAAT 300
QY 265 GlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAenGlyGluProGly 284
DB 301 GGAACCTATGGACTGTCTTTTGGTATGGAACCTCTTGTGTTTAAATGGAGAACCTGGA 360
QY 285 TyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIle 304
DB 361 TATACCATCGGACTGTCTTGTCTTTCTTTAGTGTAAATCCATAGCAGTATTGCAAT 420
QY 305 GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 324
DB 421 GGACGAGAGTCCCTCACTTTGAAACCTTCGCANTAGCCCGAGGAGCTGCTTTCATATT 480
QY 325 PheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysPro 344
DB 481 TTCCAGGTTATTGATAAACCAGGATAGATAGATAACTTTTCCACAGCTGGGATATAAACCT 540
QY 345 GluSerIleGluGlyThrValGluPheLysAenValSerPheAenTyrProSerArgPro 364
DB 541 GAATCCATAGAGGAACGTGTGGAATTTAAATATGTTCTTTCAATATATCCATCAAGACCA 600
QY 365 SerIleLysIleLeuLysGlyLeuAenLeuArgIleLysSerGlyGluThrValAlaLeu 384
DB 601 TCTATCAAGATTCTGAAGGTCTGAATCTCAGAATTAAGTCTGGAGAGACAGTCGCCTTG 660
QY 385 ValGlyLeuAenGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAsp 404
DB 661 GTCGCTCTCAATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGAT 720
QY 405 ProAspAspGlyPheIleMetValAspGluAenAspIleArgAlaLeuAenValArgHis 424
DB 721 CCGGATGATGGCTTTATCATGGTGGATGAGAAATGACATCAGAGCTTTAAATGTGCGGGAT 780
QY 425 TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSer 444
DB 781 TATCGAGACCATATTTGGAGTGGTAGTCAAGAGCTGTTTGTTCGGGACCCACCATCAGT 840
QY 445 AenAenIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArg 464
DB 841 AACCAATATCAAGTATGGACGAGATGATGCTGATGAAGAGATGCGAGAGAGCGCAAGG 900
QY 465 GluAlaAenAlaTyrAspPheIleMetGluPheProAenLysPheAenThrLeuValGly 484
DB 901 GAAGCAAAATGCGTATGATTTTATCATGAGGTTTCTTAATAATTTAATACATTTGTGAGG 960
QY 485 GluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu 504

DB 961 GAAAAAGGAGCTCAAAATGATGGAGGCGAAAAACAGAGATCGCAATTCCTGTGCTT 1020
QY 505 ValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 524
DB 1021 GTTCGAACCCCAAGATTCTGATTTTAGATGAGGCTACGCTGCCCTGGATTTCAGAAAGC 1080
QY 525 LysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValVal 544
DB 1081 AAGTCAGCTGTTCAAGCTGCACCTGGAGAGGCGAGCAAGGTCGGACTACAAATCGTGGTA 1140
QY 545 AlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMet 564
DB 1141 GCACACCGACTTCTACTATTTCGAAGTGCAGATTGATTGTGACCTTAAAGGATGGAATG 1200
QY 565 LeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeu 584
DB 1201 CTGGCGAGAAAGAGACACATGCTGAACCTAATGCGAAACAGAGTCTATATTATTCATT 1260
QY 585 ValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThr 604
DB 1261 GTGATGTGCAGGATATTTAAAAAAGCTGATGAACAGATGGAGTCAATGACATATTCTACT 1320
QY 605 GluArgLysThrAenSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle 624
DB 1321 GAAAGAAAGACCAACTCACTTCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTCAAT 1380
QY 625 AspLysAlaGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeu 644
DB 1381 GACAAGCTCAGGAATCCACCAACTTAAAGAGATAGTCTTCTGAAAGTCTCTCTATT 1440
QY 645 LysIleLeuLysLeuAenLysProGluTrpPheValValLeuGlyThrLeuAlaSer 664
DB 1441 AAAATTTAAAGTTAAACAAGCTGAATGSCCTTTGTGGTCTCTGGGACATTCGCTTCT 1500
QY 665 ValLeuAenGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMet 684
DB 1501 GTTCTAAATGGAACCTGTTCATCCAGTATTTTCCATCATCTTTGCCAAAATTTATACCATG 1560
QY 685 PheGlyAsnAenAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePhe 704
DB 1561 TTTGGAATTAATGATAAAACCAATTAAGCATGATGACAGAAATTTATTCATGATATTC 1620
QY 705 ValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArg 724
DB 1621 GTCATTTTGGTGTATTGCTTTGTCAGTTATTTTCATGAGGATTTATTTACGGCAGA 1680
QY 725 AlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGln 744
DB 1681 GCAGGGGAAATTTTAAAGATGAGATTAAAGACACTTGGCCTTCAAGGCTGTTATATCAG 1740
QY 745 AspIleAlaTrpPheAspGluLysGluAenSerThrGlyGlyLeuThrThrIleLeuAla 764
DB 1741 GATATTGCTGGTTTGTATGAAGAAAGAAAACAGCACAGAGCTTGACACAAATATTAGCC 1800
QY 765 IleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAen 784
DB 1801 ATAGATATAGCAAAATTTCAAGGAGCAACAGGTTCCAGGATTTGGCGTCTTAAACAAAA 1860
QY 785 AlaThrAenMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPhe 804
DB 1861 GCAACTAACATGGGACTTTCAGTTATCATTTCTTTATATATGATGGAGATGACATTC 1920
QY 805 LeuIleLeuSerIleAlaProValIleAlaValThrGlyMetIleGluThrAlaAlaMet 824
DB 1921 CTGATTTCTGAGTATTTGCTCCAGTACTTGGCGTGCAGAAATGATTAACCCCGCAATG 1980
QY 825 ThrGlyPheAlaAenLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGlu 844
DB 1981 ACTGATTTGCCAACAAGATAAGCAAGAACTTAAGCATGCTGGAAGATAGCAACTGAA 2040
QY 845 AlaLeuGluAenIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMet 864

Db 2041 GCTTTGAGAAATATACGTACTATAGTGTCTATTAACAAGGAAAAAGCCTTCGAGCAAAATG 2100
 QY 865 TyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGly 884
 Db 2101 TATGAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAGCACAGATATTGGA 2160
 QY 885 SerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPhe 904
 Db 2161 AGCTGTATTGCATTACGCTTATATATATTTTGGCTTATGTCAGCAGGGTTTCGATTT 2220
 QY 905 GlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThr 924
 Db 2221 GAGGCTATTTAATTCAGCTGAGCAAGATGACCCAGAGGGCCATGTTTCAATGTTTACT 2280
 QY 925 AlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSer 944
 Db 2281 GCAATTGCATATGAGCTATGGCCATCGGAAAAACGCTCGTTTGGCTCCTGAAATATCC 2340
 QY 945 LysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAsp 964
 Db 2341 AAAGCCAAATCGGGGCTCGGCATCTGTTTGGCTTGTGGAAAAAGAAACCAATATAGAC 2400
 QY 965 SerArgSerGlnGluGlyLysLysProAspThrCysGluGlyValAsnLeuGluPheArgGlu 984
 Db 2401 AGCCGAGCTCAGAGAGGAAAGAACCCAGACACATGTGAAGGAAATTTAGAGTTTCAGAA 2460
 QY 985 ValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSer 1004
 Db 2461 GTCTCTTTCTCTATCCATGTCGCCAGATGTTTTCATCTCGTGGCTTATCCCTCAGT 2520
 QY 1005 IleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSer 1024
 Db 2521 ATTAGCGAGGAAAGACAGTAGCATTTGTGGGAGCAGCGGCTGTGGGAAAGCACTTCT 2580
 QY 1025 ValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyVal 1044
 Db 2581 GTTCAACTTCTCGAGAGACTTTATGACCCCGTGCAAGGCAAGTGTCTGTTGATGTGTG 2640
 QY 1045 AspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGlu 1064
 Db 2641 GATGCAAAAGAAATTGAATGTACAGTGGCTCCGTTCCCAATAGCANCTGTTCTCAAGAG 2700
 QY 1065 ProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgVal 1084
 Db 2701 CCGTGTCTCTTCAACTGCAGCATTTGCTGAGAACATCGCCTATGTCACAAACAGCCGTGTG 2760
 QY 1085 ValProLeuAspGluIleLysGluAlaAlaAsnAlaAsnIleHisSerPheIleGlu 1104
 Db 2761 GTGCCATTAGATGAGATCAAGAAGCCGCAAAATGCAGCAAAATCCATTCTTTTATGAA 2820
 QY 1105 GlyLeuProGluLysTyrAsnThrGlnValGlyLysGlyAlaGlnLeuSerGlyGly 1124
 Db 2821 GGTCTCCCTGAGAAATACAAACACACAGTTGGACTGAAGGAGCACAGCTTTCGGCGGC 2880
 QY 1125 GlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeu 1144
 Db 2881 CAGAAACAAAGACTAGCTATTGCAAGGCTCTTCTCCAAAAACCCAAAAATTTTATGTTG 2940
 QY 1145 AspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValGlnHisAlaLeuAsp 1164
 Db 2941 GATGAGGCCACTTCAGCCCTCGCAATATGACGTGAAGAGTGGTTTCAGATGCCCTTGTAT 3000
 QY 1165 LysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsn 1184
 Db 3001 AAAGCCAGAGCGGGAAGGACATGCTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAAC 3060
 QY 1185 AlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlu 1204
 Db 3061 GCAGATTTCATAGTGTCTTGCACAATGGAAGATAAAGGAAACCAAGAACTCATCAAGAG 3120
 QY 1205 LeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 1222
 Db 3121 CTCCTGAGAAATCGACATATATATTTTAAGTTAGTGAATGCACAGTCAGTCGAC 3174

RESULT 5
 AX339028
 LOCUS AX339028 2856 bp DNA linear PAT 09-JAN-2002
 DEFINITION Sequence 10 from Patent WO0194400.
 ACCESSION AX339028
 VERSION AX339028.1 GI:18129120
 KEYWORDS human.
 SOURCE
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1
 AUTHORS Frank,M.H. and Sayegh,M.H.
 TITLE A gene encoding a multidrug resistance human p-glycoprotein
 homologue on chromosome 7p15-21 and uses thereof
 JOURNAL Patent: WO 0194400-A 10 13-DEC-2001;
 THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
 FEATURES
 source
 1..2856
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 888 a 540 c 652 g 776 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 2,27e-296 Length: 2856
 Score: 4561.50 Matches: 912
 Percent Similarity: 97.96% Conservative: 0
 Best Local Similarity: 97.96% Mismatches: 1
 Query Match: 74.06% Indels: 18
 DB: 6 Gaps: 1
 US-09-873-409-5 (1-1222) x AX339028 (1-2856)
 QY 293 ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHis-PheG1 312
 Db 114 GTTTCCTTTAGTGAATCCATAGCAGTTATTGCTATGGAGCAGCAGTCCCTCATTTATT-- 171
 QY 312 uThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPr 332
 Db 172 -----GATAAGAAACC 182
 QY 332 oSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValG1 352
 Db 183 CAGTATAGATAACTTTTCCACAGCTGGATATAAACCCTGAATCCATAGAGGAACCTGTGA 242
 QY 352 uPheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLe 372
 Db 243 ATTTAAAAATGTTCTTTTCAATTATCCATCAAGACCATCTATCAAGATTTCTGAAAGGTCT 302
 QY 372 uAsnLeuArgIleLysSerGlyThrValAlaLeuValGlyLeuAsnGlySerGlyLys 392
 Db 303 GAATCTCAGAAATTAAGTCTGGAGACAGCTCCCTTGGTCGGTCTCAATGGCAGTGGAA 362
 QY 392 sSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVa 412
 Db 363 GAGTACGGTAGTCCAGCTTCTCGCAGAGGTTATATGATCCGATGATGCTTTATCATGT 422
 QY 412 lAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVa 432
 Db 423 GGATGAGAATGCATCAGAGCTTTAAATGTGCGGATTTATTCGAGACCATATTGGAGTGT 482
 QY 432 lSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAs 452
 Db 483 TAGTCAAGAGCCTGTTTGTTCGGGACCACCATCAGTAAACAATATCAAGTATGGACGAGA 542
 QY 452 pAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheI1 472
 Db 543 TGATGTGACTGATGAAGATGAGAGAGCAGCAGGAGGAGCAAAATCGTATGATTTTAT 602
 QY 472 eMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerG1 492

Db 603 CATGGAGTTCCTAATAAATTTAATCATTTGGTAGGGAAAAAGAGCTCAAAATGAGTGG 662
QY 492 YGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProlysisIleLeuI 512
Db 663 AGGCAGAAACAGAGATCGCAATTCCTGCTGCTTAGTTCGAAACCCCAAGATTCTGAT 722
QY 512 eLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaLe 532
Db 723 TTTAGATGAGGCTACGCTCGCTCGCTGATTTCAGAAAGCAAGTCAGCTGTTCAAGCTGCACT 782
QY 532 uGluLysAlaSerLysGlyArgThrThrIleValValAlaAlaHisArgLeuSerThrIleAr 552
Db 783 GGAGAGCGCAGCAAGGTCGAGTACAAATCGTGGTAGCACCGCACTTTCTACTATTTC 842
QY 552 gSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAl 572
Db 843 AAGTGCAGATTGATTGTGACCCCTAAAGGATGGATGCTGGCGAGAAAGAGACACATGC 902
QY 572 aGluLeuMetAlaLysArgGlyLeuTyTySerLeuValMetSerGlnAspIleLysLy 592
Db 903 TGAACATAATGGCAAAACAGAGTCTATATTATTCATTGTGATGTCACAGGATATTAATA 962
QY 592 sAlaAspGluGlnMetGlnSerMetThrTySerThrGluArgLysThrAsnSerLeuPr 612
Db 963 AGCTGATGAACAGATGGAGTCAATGACATATTCCTGAAAGAAAGACCAACTCACTTCC 1022
QY 612 oLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluSerThrGl 632
Db 1023 TCTGCATCTGTGNAAGAGCATCAAGTCAGACTTCATTGACACAGCTGAGGATCCACCCA 1082
QY 632 nSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysPr 652
Db 1083 ATCTAAAGAGATAGTCTCTCTGAAGTCTCTCTATTATAAATAATTTAAAGTTAAACAAGCC 1142
QY 652 oGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPr 672
Db 1143 TGAATGGCCCTTTGTGTGTTCTGGGGACATTTGGCTTCTGTCTAAATGGAACTGTTCATCC 1202
QY 672 oValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrTh 692
Db 1203 AGTATTTTCATCATCTTTGCAAAAATTAATACCATGTTTGGAAATTAATGATTAACACCAC 1262
QY 692 rLeuLysHisAspAlaGluIleTySerMetIlePheValIleLeuGlyValIleCysPh 712
Db 1263 ATTAAGCATGATGCAGAAATTTATTCATGATATTCGTCAATTTGGTGTGTTATTCCTT 1322
QY 712 eValSerTyPheMetGlnGlyLeuPheTyGlyArgAlaGlyGluIleLeuThrMetAr 732
Db 1323 TGTCAATTTATTCATGCGGATTTATTTACGGCAGAGCAGGGGAAATTTTAACGATGAG 1382
QY 732 gLeuArgHisLeuAlaPheLysAlaMetLeuTyGlnAspIleAlaTrpPheAspGluLy 752
Db 1383 ATTAAGACACTTGGCCCTTCAAGCCATGTTATATACAGATATTCCTCGTTTGGATTGAAAA 1442
QY 752 sGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGl 772
Db 1443 GGAACACAGCACAGAGGCTTGACAACTATATTAGCCATAGATATAGCAAAATTCAGG 1502
QY 772 yAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerVa 792
Db 1503 AGCAACAGGTTCCAGGATTTGGCGTCTTTAAACACAAATGCAACTAACATGGGACTTTTCAGT 1562
QY 792 lIleIleSerPheIleTyGlyTyTrpGluMetThrPheLeuIleLeuSerIleAlaProVa 812
Db 1563 TATCATTTCTTATATATGATGGAGATGACATTCCTGATTCAGATATTCGATTCAGT 1622
QY 812 lLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLy 832
Db 1623 ACTTGGCTGTACAGGAATGATTGAACCGCAGCAATGACTGATTTTGCCAAACAAAGATAA 1682
QY 832 sGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIl 852
Db 1683 GCAAGAACTTAAGCATGCTGGAAAGATAGCAACTGAAGCTTTTGAGAAATATACGTACTAT 1742

QY 852 eValSerLeuThrArgGluLysAlaPheGluGlnMetTyGluGluMetLeuGlnThrGl 872
Db 1743 AGTGTCAATTAACAAGGCAAAAGCCCTTCAGCAAAATGATGAAGAGATGCTTCAGACTCA 1802
QY 872 nHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyAlaPheSerHisAl 892
Db 1803 ACACAGAAATACCTCGAAGAAAGCACAGATATTGGAAGCTGTATTGCAATTCAGCATGC 1862
QY 892 aPheIleTyPheAlaTyAlaAlaGlyPheArgPheGlyAlaTyLeuIleGlnAlaGl 912
Db 1863 CTTTATATATTTTGCCTATGCAGAGGTTTCGATTTGAGGCTATTATTAATTCAGACTG 1922
QY 912 yArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyGlyAlaMetAl 932
Db 1923 ACGAATGACCCAGAGGCGCATGTTTCATAGTCTTTTACTGCAATTCGATATGGAGCTATGGC 1982
QY 932 alleGlyLysThrLeuValLeuAlaProGluTySerLysAlaLysSerGlyAlaAlaHi 952
Db 1983 CATCGGAAAAACGCTCGTTTGGCTCTCTGAATATTCCAAAGCCAAATCGGGGCTCGCA 2042
QY 952 sLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLy 972
Db 2043 TCTGTTTGGCTTGTGGAAGAAACCAATATAGACAGCCGAGTCAAGAGGGAANA 2102
QY 972 sProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyProCysAr 992
Db 2103 GCCAGACACATGTGAAGGAAATTTAGAGTTTCGAGAAGTCTCTTCTCTATCATGTCG 2162
QY 992 gProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAl 1012
Db 2163 CCAGATGTTTTCATCCTCGCTGCTTATCCCTCAGTATTGAGCGAGGAAAGACAGTAGC 2222
QY 1012 aPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTy 1032
Db 2223 ATTTTGGGGAGCAGCGCTGTGGAAAGACACTTCTGTTCACTTCTGCAGAGACTTTA 2282
QY 1032 rAspProValGlnGlyValLeuPheAspGlyValAspAlaLysGluLeuAsnValGl 1052
Db 2283 TGACCCCGTGCAAGGCAAGTGTCTGTTGATGTGTGATGCAAAAGAAATTAATGATGTA 2342
QY 1052 nTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIl 1072
Db 2343 GTGGCTCCGTTCCAAATAGCAATCGTTCCTCAAGAGCTGTGCTCTTCACTTCACTGCA 2402
QY 1072 eAlaGluAsnIleAlaTyGlyAspAsnSerArgValValProLeuAspGluIleLysGl 1092
Db 2403 TGTGAGAACATCGCTATGTTGACCAACAGCCGTGTGTGCCATTAGATGAGATCAAGA 2462
QY 1092 uAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyArgnth 1112
Db 2463 AGCCGCAATGCAGCAAAATATCCATTTCTTTATTGAAGGTCTCCCTGAGAAATACAACAC 2522
QY 1112 rGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAl 1132
Db 2523 ACAAGTTGGACTGAAAGGAGCACAGCTTTCTGGCGGCCAGAAAAACAAAGACTAGCTATTGC 2582
QY 1132 aArgAlaLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAs 1152
Db 2583 AAGGGCTCTTCTCCAAAAACCCAAATTTTATTGTTGATGAGGCCACTTCACGCCCTCGA 2642
QY 1152 pAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCy 1172
Db 2643 TAATGACAGTGAGAGGTGTTTCAGCATGCCCTTGATAAAGCCAGGACGGAGGACATG 2702
QY 1172 sLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHi 1192
Db 2703 CCTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAACGCGCAGATTGTAGTGTGTTCTGCA 2762
QY 1192 sAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTy 1212
Db 2763 CAATGGAAAGATAAAGAACACAGAACTCATCAAGAGCTCTCGAGAAATCGAGACATATA 2822


```

QY 1212 rPheLeuValAsnAlaGlnSerValGln 1222
|||||
Db 2823 TTTTAAAGTAGTGAATGCACAGTCAGTCAG 2853

RESULT 6
LOCUS CRUGPPI 4296 bp mRNA linear ROD 22-JUN-1995
DEFINITION C.griseus P-glycoprotein (isoform class I) mRNA, complete cds.
VERSION M60040.1 GI:191164
KEYWORDS multidrug resistance; p-glycoprotein; transmembrane protein.
SOURCE C.griseus adult liver and ovary, cDNA to mRNA.
ORGANISM Cricetulus griseus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
Cricetulus.
1 (bases 1 to 4296)
Endicott, J.A., Sarangi, F. and Ling, V.
Complete cDNA sequences encoding the Chinese hamster P-glycoprotein
gene family
DNA Seq. 2 (2), 89-101 (1991)
MEDLINE 92135896
PUBMED 1685679
FEATURES
    Location/Qualifiers
        1..4296
            /organism="Cricetulus griseus"
            /db_xref="taxon:10029"
            /tissue_type="liver and ovary"
            /dev_stage="adult"
            103..3933
            /note="putative"
            /codon_start=1
            /product="p-glycoprotein isoform I"
            /protein_id="AAA68883.1"
            /db_xref="GI:191165"
            /translation="MEFDEFSGRKDNFLKMGKSKKKEKPKVSVFTMPYVAGW
LDRLVNLVGLTAAIHGVALPLMLVFGDMTDSFASVGNIPFNATNNTATQNASDIFG
KLEBEEMTYAYITGIGAGLVIAVQVSWFCLAAGRQIHKIRQKFFHAIMNKGWIF
DVHDEGLNLTDDVSINIGIGDKIGMFQMATFFGGFIIGTRGKMLTLVILAI
SPVLGSLAGIWAKEISPTDKELQAYAKAGAAVEVLAAIRTVIAFGQKKELEYNN
NLEAKRLGICAKITANTSMGAAPFLIVASVALAPWYGTSLVISEYISIGOVLTVFFA
VLIGAFSIGQAPNTEAFANARGAAVEIFNIIDNKPESIDSPSKNGYKPDNIKGLFEK
NIHFSYPRKQVQIILKGLNLKQSQTVLNVGNSCGKSTTVQLLQRLYDPFEGVWSI
DQDILIRNRYLRIEILGVWSQEPVLFATTAENIRYGRNVITMDEIEKAVKEANAYD
FIMKLPKFDITLVGRGQALQSGQRIALPVRNPKILLIDATLSDATDSEAVV
QAAIDKAREGRTTIVIAHRLSTVRNADIAGPDGVIYVQGNHBEIMREKGIYFKLVN
TOTAGNEIELGNVEGESKNEIDNLDMSKDSASSLIRRSSTERSIRGPHDQDRKLSLK
EALDEDVPISPWRILKLSNSEWPFYVIGIFCAIYNGALQAPSIIFSKYVGVFRNT
DDETRKHSNLSFLFLILGVISFIFFLQGFQKAGEILTKRLRYWFKMSMLQDV
SWFDPNKTGATTTLRMDAGQVKGATCARLAVITONIANLGTGIIISLIYGMQLTL
LLLAIVPIIAIAGVVMKLSQALKDKKELESGKIATEAIENTFTVSLTRQKFE
NNYAQSLQIPYRNALKKAHVFGITFTSQAMMYFYAACFRFGAYLVARELWTPENVL
LVFSAIVFGMAAGVQSPFADPYAKAVSASHIIMIEKVPISDSYSTGGLKPNLTLEG
NVKNEVFNPTFRDIPVLQGLNLEVKQTLALVGSCKSTVQVQLLERFYDPMMA
GTVFDLGGKESLQVWMLRHGIVSQEPILFDCSIAENIAYGDNRSVQDEIERAA
KEANTHQIESLPDKNYTRVGDKGTQLSGGQKQRIAIARALVRQPHILLDEATSALD
TESKVOEALDKAREGRTTIVIAHRLSTIQNADLIIVIQNGKVKHEGTHQQLAQKG
IYFSMVSVQAGAKR"
BASE COUNT 1271 a 892 c 1029 g 1104 t
ORIGIN

Alignment Scores:
Pred. No.: 2,05e-228 Length: 4296
Score: 3550.50 Matches: 679
Percent Similarity: 73.75% Conservative: 248
Best Local Similarity: 54.02% Mismatches: 259
Query Match: 57.65% Indels: 71
DB: 10 Gaps: 9

US-09-873-409-5 (1-1222) x CRUGPPI (1-4296)

QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
|||||
Db 1228 TCAAGAAATGGGTACAAACCAGACCAACATTAAGGAAATTTGGAAATTTCAAAATATTTCAC 1287

```

```

Db 250 ATGCTGGTGGGACTCTCGCAGCTATTATCCATGGAGTTGGACTCCCGCTTATGATGCTG 309
QY 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsn----- 38
|||:::|||||:::|||||:::
Db 310 GTGTTTGGAGACATGACAGACAGCTTTGCAAGTGTAGGAAACATCCCAACAATGCTACT 369
QY 38 ----- 38
Db 370 AATAATGCCACTCAAGTCAATGCTTCAGACATCTTTGGCAAACTGGAGGAAGAAATGACC 429
QY 39 ThrTyrSerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyTyrPheAspSerCysAsp 58
|||:::|||||:::|||||:::|||||:::|||||:::
Db 430 AGTACGCCCTAT-----TACTACACAGGAGATTGGTCTGGTGGCTTATA 474
QY 59 PheGlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrIleAspGile 78
|||:::|||||:::|||||:::|||||:::|||||:::
Db 475 GTTGCCATACATTGAGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAT 534
QY 79 ArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAsp 98
|||:::|||||:::|||||:::|||||:::|||||:::
Db 535 AGACAGAAAGTTTTTTTCATGCTATAATGAATCAGAGATAGGCTGGTTTACGCTGCACGAC 594
QY 99 IleGlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGly 117
|||:::|||||:::|||||:::|||||:::|||||:::
Db 595 GTTGGGAGCTCAACACCCGGCTCACAGATGATGCTCCAAATTAACGAGGAATGTG 654
QY 118 AspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGly 137
|||:::|||||:::|||||:::|||||:::|||||:::
Db 655 GACAAATTTGGAATGTTCTTTCAGGCAATGGCAACATTTTTTGGTGGTTTATAATAGGA 714
QY 138 LeuValLysGlyTyrIleLeuThrLeuValThrLeuSerThrSerProLeuIleMetAla 157
|||:::|||||:::|||||:::|||||:::|||||:::
Db 715 TTTTACTCGAGGCTGGAAGCTAACCTTGTGATTTTGGCCATCAGCCCTGTTCTGGACTG 774
QY 158 SerAlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyr 177
|||:::|||||:::|||||:::|||||:::|||||:::
Db 775 TCAGCTGGTATTGGGCAAGATATATCTTCTATTACTGATAAGAAGCTTCAGGCATAT 834
QY 178 SerLysAlaGlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPhe 197
|||:::|||||:::|||||:::|||||:::|||||:::
Db 835 GCAAAAGCTGGAGCAGTGTCTCAAGAAGTCTTAGCAGCCATCAGAACTGTGATGCTATT 894
QY 198 ArgAlaGlnGlyGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrp 217
|||:::|||||:::|||||:::|||||:::|||||:::
Db 895 CGAGACAAAAGAAAGAACTTGA----- 918
QY 218 PheTyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsn 237
|||:::|||||:::|||||:::|||||:::|||||:::
Db 919 -----AGGTACACACAAAT 933
QY 238 LeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGly 257
|||:::|||||:::|||||:::|||||:::|||||:::
Db 934 TTAGAAGAGCCAAAAGACTTGGGATAAAGAAAGCGATCACAGCCACATTTCCATGGGC 993
QY 258 AlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrGlyThrSerLeu 277
|||:::|||||:::|||||:::|||||:::|||||:::
Db 994 GCAGCTTTCTTCTTATCTATCATCATCATATGCTCTGGCCTTCTGGTATGGAACGCTCTTG 1053
QY 278 IleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPheSerVal 297
|||:::|||||:::|||||:::|||||:::|||||:::
Db 1054 GTCATCTCAAAAGAA-----TATTCTATTGGCAAGTCTCACTGTCTTCTTGTGTA 1107
QY 298 IleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAla 317
|||:::|||||:::|||||:::|||||:::|||||:::
Db 1108 TTAATTGGGCATTCAGTATTGGACAGGATCTCCAAATATCGAAGCCCTTTCATGCA 1167
QY 318 ArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPhe 337
|||:::|||||:::|||||:::|||||:::|||||:::
Db 1168 AGAGGAGCAGCCTATGAAATCTTCAACATAATTGATAAAGCCCAAGTATTGACAGCTTC 1227
QY 338 SerThrAlaGlyTyrIleProGluSerIleGluGlyThrValGluPheLysAsnValSer 357
|||:::|||||:::|||||:::|||||:::|||||:::
Db 1228 TCAAGAAATGGGTACAAACCAGACCAACATTAAGGAAATTTGGAAATTTCAAAATATTTCAC 1287

```


Qy	358	PheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLys	377
Db	1288	TTcAGTTACCCATCTCGAAAGACGTTcAGATCTTTGAAGGCGCTCAACCTGAAGGTGCAG	1347
Qy	378	SerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGln	397
Db	1348	AGCGGACAGACAGTGGCCCTGGTTGGCAACAGCTGCTGTGGGAAAGACCACTGTCCAG	1407
Qy	398	LeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIle	417
Db	1408	CTGCTGCAGAGGCTTcAGNACCCACAGAGGCGGTGCTcAGTATCGACGGTCAGACATC	1467
Qy	418	ArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProVal	437
Db	1468	AGGACCATCAATGTGAGGTATCTCGCGGAAATCAITTTGGGTCGTGAGTCAGGAACCTGTG	1527
Qy	438	LeupheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGlu	457
Db	1528	TTGTTTCCCACTcATCGCTGAAACATTTGCTCTATGCGCCGAGAAAATGTcACATATCGAC	1587
Qy	458	GluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsn	477
Db	1588	GAGATCGAGAAGCTGTCAAGGAGCCATGCTATGACTTCATCATGAACCTGCCCAT	1647
Qy	478	LysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArg	497
Db	1648	AAATTTGCACATCTCTGTGTCAGACAGGACACAGCTcAGTGGAGGACAGAAACAGAGA	1707
Qy	498	IleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThr	517
Db	1708	ATGCCCATTTGCTGTCGCCCTGGTCCCAACCCACAGATCTTTTGGATGAGGCGACA	1767
Qy	518	SerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLys	537
Db	1768	TCACCTTTGGACACAGAAAGTGAAGCGGTGTTcAGGCTGCTCTGGATAAGSCTAGAGAA	1827
Qy	538	GlyArgThrThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIle	557
Db	1828	GGCCGGACTACCATTTGTGATAGCCACCGCTGTCTACAGTTTCGAAATTCGTGACATCAT	1887
Qy	558	ValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLys	577
Db	1888	GCTGGGTTGATGGTGTGTGTCATTGTGGAGCAGGAATCATGAAGAGCTCATGAGAG	1947
Qy	578	ArgGlyLeuTyrTyrSerLeuValMetSerGln-----	588
Db	1948	AAGGCATTTTACTTCAAACTTGTcATGACACAGCAGCAGGAATAATGAATTTGAATTAGGA	2007
Qy	589	---AspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLys	607
Db	2008	AATGAAGTTGGTGAGTCTAAAAATGAATTTGATAATTAGACATGCTCTTCAAAAAGATTCA	2067
Qy	608	ThrAsnSerLeu-----ProLeuHisSerValLysSerIleLysSerAspPheIleAsp	625
Db	2068	GCATCCAGTCTAATCAGAAGGAGATCAACTCGCAGNAGTATCCGTTGGACCATGACCA	2127
Qy	626	LysAlaGluGluSerThr---GlnSerLysGluIleSerLeuProGluValSerLeuLeu	644
Db	2128	GACAGGAAGCTTAGTACCAAGAGGCGCTTGGATGAAGATGATACCTCCAATTTCTTTTGG	2187
Qy	645	LysIleIleLysLeuAsnLysProGluTyrProPheValValLeuGlyThrLeuAlaSer	664
Db	2188	CGTATCTAAAGTTGAATTCATCTGAATGGCCATATTTTGTGGTTGGTATATTTCTGTGCC	2247
Qy	665	ValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMet	684
Db	2248	ATAGTAATGGAGCTTGCAACCAACGATCTCTCAATAATATTTTCTAAGGTTGTAGGGTT	2307
Qy	685	Phe---GlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIle	703
Db	2308	TTCAAGAAATCTGATGATGAACCAACACACATGATAGCACTTGTTCTCTTATTG	2367

Qy	704	PheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGly	723
Db	2368	TTTCTGATCCTTGGGTCATTTCTTTATTATCATATTTTCTTCAGGGCTTCACATATTTGGC	2427
Qy	724	ArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheIysAlaMetLeuTyr	743
Db	2428	AAAGCTGGAGAGATCTCTCAACCAAGGAGCTCCGATACATATGGTTTCAAATCCATGCTGAGA	2487
Qy	744	GlnAspIleAlaTrpPheAspGluIysGluAenSerThrGlyGlyLeuThrThrIleLeu	763
Db	2488	CAGGATGTGAGCTGGTTTGATTAACCTTAACCAACACCATCTGAGGACATTGACCACACGAGCTC	2547
Qy	764	AlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGln	783
Db	2548	GCCAAACGATCTGGTCAAGTTAAAGGGGCTACAGGAGCAGACTGCTGTCTCATTTACCCAG	2607
Qy	784	AsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyrPheMetThr	803
Db	2608	AATATAGCAAAATCTTGGGACAGGAATCATCATATCCCTTAATCTATGGCTGGCAGTTGACA	2667
Qy	804	PheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAla	823
Db	2668	CTTCTACTCTTAGCAATTTGTCCTCCATCATTCGATACAGAGAGTGTGAGATGAAAATG	2727
Qy	824	MetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThr	843
Db	2728	TTGTCTGGACAAGCACTAAAGATTAAGAGGAGCTAGAAAGTTCTTGGGAAGATTTGCTACT	2787
Qy	844	GluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGln	863
Db	2788	GAAGCAATAGAGAACTTCCGCATCTGCTCTTTGACTCCGGAGCAGAACTTTGAAAAAT	2847
Qy	864	MetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIle	883
Db	2848	ATGTATCCCGAGAGCTTGCAGATACCTACAGAAATGCTCTGAAGAAGACACAGCTCTTT	2907
Qy	884	GlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArg	903
Db	2908	GGGATCACATTTCTCCTTCAACCCAGGCCATGATGTTATTTTCGTATGCTCTGTTTCCGG	2967
Qy	904	PheGlyAlaTyrIleuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPhe	923
Db	2968	TTTGGTCTTACTTGGTGGCAGCGGAATTAATGACATTTGAAATGTTCTATTAGTATTC	3027
Qy	924	ThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValIleAlaProGluTyr	943
Db	3028	TCAGCTATTGTTCTGGTCGATGGCAGCTGGCGAGGTGCTGTTCAATTTGCTCTCTGACAT	3087
Qy	944	SerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGlyLysLysProAsnIle	963
Db	3088	GCCAAAGCCAAAGTGTGGCATCGCATATCATCATGATTCATTGAAAAAGTCCCTTCCATT	3147
Qy	964	AspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArg	983
Db	3148	GACAGCTACACGACCGAGCGCTTGAAGCCCTAATACATATGGAAGGAAATGTGAAATTTAAT	3207
Qy	984	GluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeu	1003
Db	3208	GAAGTCGTGTTCAACTACCCACCCCGGCTTGACATCCAGCATGCTTCAGGGGCTGACCTG	3267
Qy	1004	SerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThr	1023
Db	3268	GAGGTGAAGAAGGGCCAGACGCTGGCCCTTGGTGGGCAGCAGTGGCTGCGGGAAGAGCACA	3327
Qy	1024	SerValGlnIleuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGly	1043
Db	3328	GTGCTCCAGCTGCTCGAGCGGTTCTACGACCCCATGGCGGGAACAGTGTCTTAGATGGC	3387
Qy	1044	ValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGln	1063
Db	3388	AAAAGAGTAACCACTAAATGTCCAGTGGCTCCGAGACACCATGAGCATTTGTGTCCCA	3447
Qy	1064	GluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArg	1083


```
QY 924 ThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyr 943
Db 3036 TCAGCTATTGCTGTCGTCATGGCAGTGGCAGGTTCAGTTTCATTTGCTGCTGACT 3095
QY 944 SerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIle 963
Db 3096 GCCAAAGCCAAAGTGCAGCATCGCATATCATCATCATGTAAGAAAGTCCCTTCATTT 3155
QY 964 AspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArg 983
Db 3156 GACAGCTACAGCACCGGAGCCCTGAAGCCTATACATTGGAAGGAATGTGAATTAAT 3215
QY 984 GluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeu 1003
Db 3216 GAAGTCGTGTTCAACTACCCACCGCCCTGCATCCAGTGCCTTCAGGGGCTGAACCTG 3275
QY 1004 SerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThr 1023
Db 3276 GAGGTGAAGAAGGCGCAGACGCTGGCCCTGGTGGGCGAGCAGTGGCTGCGGGAAGACCA 3335
QY 1024 SerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGly 1043
Db 3336 GTGTCCTCAGCTGCTCAGCGGTTCTACGACCCCATCGCCGGAACAGTGTTCCTAGATGC 3395
QY 1044 ValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGln 1063
Db 3396 AAAGAAGTAACCAACTAAATGTCCAGTGGCTCCGAGCACACCTAGGCATTTGTGCCAA 3455
QY 1064 GluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArg 1083
Db 3456 GAGCCCATCTGTTGACTGCAGCATCGCTGGAACAATCGCTTATGGAGACAAACGCCGG 3515
QY 1084 ValValProLeuAspGluIleLysGluAlaAlaAsnAlaAsnIleHisSerPheIle 1103
Db 3516 GTGCTGCCAGTCCAGATCGATCGAGAGCGGCGGCAAGGAGCAACATCCACCACTTCA 3575
QY 1104 GluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGly 1123
Db 3576 GAGTCACCTGCTCAGCAATACACACACAGTAGGAGACAAAGGAGTCTCAGCTGTCGGGT 3635
QY 1124 GlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeu 1143
Db 3636 GGGCAGAAACACGTCATCGCATAGCAGCGCCCTCGTCAGACAGCCTCACAATTTTGCTC 3695
QY 1144 LeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeu 1163
Db 3696 CTGGATGAAGCAACATCAGCCCTGGATACAGAGATGAAGATTTGTCAGGAGCTCTG 3755
QY 1164 AspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGln 1183
Db 3756 GACAAAGCCAGAGAGGCGCACCTGTCATTGTGATCGCTCACCGCTGTCCACCATCCAG 3815
QY 1184 AsnAlaAspIleValValLeuHisAsnGlyValIleLysGluGlnGlyThrHisGln 1203
Db 3816 AACCGAGACTGATCGTGGTATTCAGATGTCAGAGGTCAAGAGGATGCGACCCACCAG 3875
QY 1204 GluLeuLeuArgAsnAspIleTyrPheLysLeuValAlaGlnSer 1220
Db 3876 CAGCTGCTGCACAGAAAGCATATTTCTCCATGCTAGTGTGAGGCT 3926
```

```
RESULT 8
LOCUS AX105082 4279 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 26 from Patent WO0123540.
ACCESSION AX105082
VERSION AX105082.1 GI:13921232
KEYWORDS dog.
SOURCE Canis familiaris
ORGANISM Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE 1 (bases 1 to 4279)
AUTHORS Stocker,P.J., Steimel-Crespi,D.T., Crespi,C.L., Reif,T.C. and
```

```
Patten,C.J.
P-glycoproteins and uses thereof
Patent: WO 0123540-A 26 05-APR-2001;
GENTEST CORPORATION (US)
FEATURES
Location/Qualifiers
source 1..4279
/organism="Canis familiaris"
/db_xref="taxon:9615"
17..3862
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC37723.1"
/db_xref="GI:13921233"
/translation="MDPEGRKGSAAENKFWMGKSKKKEKKKPTVSTFAMFRYSN
WLDRLYMLVGTMAAI IHGAALPLMWLFVGNMNTDSFANAGI SRNKTFPI INESITNT
EIGHNLEEMETVAYVYGGVAGVLAAYVQVFWCLAGKQILKIRKQFIMRQ
QHFVDHVDGELNRLTDDVSKNEGIDKIGMFFQSIATFTFTGFI VGFTRGWLTL
VILAI SPVLGSAI WAKILSSFTDKELLAYAKAGAAVEVLAIRTVAEGGQKEL
EYNNLEEAAGIGIKKAITANISIGAAFLLIYASIALAFWGTSLVSSEYTGQVL
TVFVSLIGAFSIGQASPIEAFANAGAAYEIKIDNKPISIDSYSGHKPDNIKG
NLEFNHFSYPSRKEVKILKGLNKVQSGQTVLVNCGGCKSTTVQLMQRLYDPTD
GMVICDQDIRTINVRLREITGVVSOBVLPAITAEINIRYGRNVTMDIEKAVKE
ANAYDFIMKLPNKEDTLVGERGAOLSGGQKORIAIARALVRNPKILLDEATSLDTE
SEAVVQALDKARKGRTTIVIAHRLSTVRNADVIAGPDGVLVEKGNHDELMKEGII
FKLVMTQTRNEIELENAIGESKESDALESPKDSGSLIKRRSTRKSIHAPQGDQR
KLTGKEDLNENVPVPSFMRILKLNSTEPYFVVGIFCAIINGGLOPAFSIIPSRIGI
PTRDEDPEKTRQNSMFSVLVLGIIISFTIFLQGTFTGKAGEILTKRLRYMFRSM
LRQVDFDPDKNTTALTTLANDAAVGAIGSLRAVITONTIANLGTIIISIIYR
WOLTLVLLAI VPIIAIAGVEMKLSGOALKDKKELEGAKIATEIENFRTVLSLR
EQKEEYMAQSLQVPRNLSKAHIFGVSFISITQNMVYFSGACFRGAYLVANFPMN
FDVLLVFSALVFGAMAVQSSFAPDIYAKARVSAHVIMIEKSPIDSYSYSGHGLKP
NTLEGNVTNEVFNYPTRDIPVLQSLVSKVKGQTLALVSGSCGSKSTVVQLLERF
YDLAAGSVLIDGKEIKHLNWLRAHLGIVSQEPLFDSCIAENIAYGDNRSRVSHEE
IVQAAKEANIHHFTETLPEKYNTRVGDGTOLSGGQKORIAIARALVRNPKILLDEA
TSALDTESEKVQALDKAREGRTICIVIAHRLSTIQNADLIVVFQNGKVEHGTQQQL
LAQKGFPSVMVSVQAGKR"
BASE COUNT 1296 a 833 c 1009 g 1141 t
ORIGIN
Alignment Scores:
Pred. No.: 2,8e-227 Length: 4279
Score: 3533.50 Matches: 683
Percent Similarity: 73.13% Conservative: 237
Best Local Similarity: 54.29% Mismatches: 269
Query Match: 57.37% Indels: 69
DB: 6 Gaps: 9
US-09-873-409-5 (1-1222) x AX105082 (1-4279)
QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 167 ATGTTGGTGGGACAAATGCTGCCATCATCCATGGAGCTGCACCTCCCTCTCATGATGCTG 226
QY 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu----- 34
Db 227 GTTTTGGAAACATGACAGATAGCTTTGCAAAATGCAAGGAATTTCAAGAAACAAACTTTT 286
QY 35 -----ValGlnThrAsnThrTyrSerPhePheArg----- 44
Db 287 CCAGTTATAATTAAGAAAGTATTAGCAACAATACACAACATTTTCATCAACCTCTGGAG 346
QY 45 -----LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe 59
Db 347 GAGGAATGACCAACGATGCTTATTATACAGTGGGATCGGTGCTGCTGCTGGTGGCT 406
QY 60 GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg 79
Db 407 GCTTACATCCAGGTTTCATCTGCTGCTGCGCAGGAGGAGACAGATACTCAAAATTAGA 466
QY 80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99
Db 467 AAACAATTTTTTTCATGCTATCATGCGACAGGAGATGGCTGGTTTACGTCGATGACGTT 526
```


QY 100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
DB 527 GGGGAGCTTAACACCGGCTCACAGACGATGTCTCCAAAATCAATGAAGAAATGGCGAC 586
QY 119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
DB 587 AAAATTGGAAATGTTCTTCAATCAATAGCAACATTTTTCACCGGTTTATAGTGGGGTTT 646
QY 139 ValIleGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
DB 647 ACAGGTGGTGGAGACTACCTTGTGATTTTGGCCATCAGCCCTGTCTTGACATTTCA 706
QY 159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaIleThrSer 178
DB 707 GCCGCCATCTGGCAAGATACCTATCTTCTTACTGATTAAGAACTCTTGGCCTATGCA 766
QY 179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198
DB 767 AAAGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGAACTGTGATTTGCCTTTGA 826
QY 199 AlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPhe 218
DB 827 GGCAAAAGAAAGAACTTGAA----- 847
QY 219 TyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeu 238
DB 848 -----AGGTACAACAAAAATTTA 865
QY 239 LysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAla 258
DB 866 GAAGAAGCTAAAGGAATTTGGGATAAAGAAAGCTATCAAGCCCAACATTTCTATTGTGGCC 925
QY 259 ValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIle 278
DB 926 GCTTCTTATGATCATGATCATATGCTCTGGCTTCTGGTATGGGACTCTTGGTC 985
QY 279 LeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIle 298
DB 986 CTCTCCAGTGAA-----TATACTATTGGACAAGTACTCACTGCTCTCTTTCTGTATTA 1039
QY 299 HisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArg 318
DB 1040 ATTTGGGGCTTTTATGATTGGACAGGCATCCCAAGCATTTGAAGCATTTGCAACGCAAGA 1099
QY 319 GlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSer 338
DB 1100 GGAGCCTCTATGAAATCTTCAGATTAATTTGACATAAACAAGCATTTGACGCTATTCG 1159
QY 339 ThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPhe 358
DB 1160 AAGAGTGGACATAAAACAGATAATATTAAAGGAAATTTGGAAATTTCAAAAATGTTCACTTC 1219
QY 359 AsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSer 378
DB 1220 AGTTACCTCTTCGAAAAGAAGTTAAGATCTTAAAGGCTCTCAACCTGGAAGGTTACAGAT 1279
QY 379 GlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeu 398
DB 1280 GGGCAGACAGTGGCGCTGGTTGGAAACAGTGGCTGGGGAAGACACACCGTGCAGCTG 1339
QY 399 LeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArg 418
DB 1340 ATGCAGAGGCTCTATGACCCACAGATGGCATGCTGTATTGATGACAGGACATTTAGG 1399
QY 419 AlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeu 438
DB 1400 ACCATAAATGTAAGGCATCTCTCGGAAATTTACTGGTGGTGGATGTCAGAGCCTGTGTTG 1459
QY 439 PheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspValThrAspGluGlu 458
DB 1460 TTTGCCACACAGATGATGAAAACATTTGCTATGTCGCCGCGAAATGTCACCATGGATGAG 1519
QY 459 MetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLys 478

DB 1520 ATTGAGAAAGCTGTTAAGGAAGCCCAATGCCTATGATTTTATCATGAACACTACTCAATAAA 1579
QY 479 PheAsnThrLeuValGlyGlyLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIle 498
DB 1580 TTTGACACTCTGTTGGAGAGAGAGGGGCCAGCTGAGTGGTGGACAGAAACAGAAATC 1639
QY 499 AlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSer 518
DB 1640 GCCATTGCTCGGCCCTGCTTCGCAACCCCAAGATTTCTTCTGCTGATGAGGCAACGTCA 1699
QY 519 AlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGly 538
DB 1700 GCTCTGACACTCAAAAGTGAAGCAGTGGTTTCAGGTGGCCCTGATAGGACCAAGAAAGGC 1759
QY 539 ArgThrThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleVal 558
DB 1760 CGGACTACCATTTGTGATAGTCATCTGTTGTCTACAGTTTCGTAATGCCGATGTCATGCT 1819
QY 559 ThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArg 578
DB 1820 GGTTTTGTATGATGGAGTCATTGTGGAGAAAGAAATCATGATGAACCTCATGAAGAGAAG 1879
QY 579 GlyLeuTyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGlu 598
DB 1880 GGCATTACTTCAAACTTGTCAATGTCAG---ACAAGAGGAAATGAAATTGAGTTAGAA 1936
QY 599 SerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSer 618
DB 1937 ATGCCACTGGTGAATCCAAAAGTGAAGTGAAGTGCCTTGGAAATG---TCTCCAAAGAT 1993
QY 619 IleLysSerAspPheIle----- 624
DB 1994 TCAGGCTCCAGTTTAATAAAAAAGAAAGATCAACTCGCAGGAGTATACATGCACCACAAGGC 2053
QY 625 ---AspLysAlaGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeu 643
DB 2054 CAAGACAGAAAGCTTGGTACAAAAGAGGACTTCAATGAGAATGTACTCCAGTTTCTTTC 2113
QY 644 LeuLysIleLeuLysAsnLysProGluTrpPropheValValLeuGlyThrLeuAla 663
DB 2114 TGGAGATTTCTGAAGCTGAATCACTCACTGAATGGCTTATTTGTGGTGGTATATTTGT 2173
QY 664 SerValLeuAsnGlyThrValHisProValPheSerIlePheAlaLysIleIleThr 683
DB 2174 GCTATTATAACCGAGGCTGCAACCCAGCATTTTCAATAATATTTTCAAGGATTAAGGG 2233
QY 684 MetPhe---GlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMet 702
DB 2234 ATCTTTACCCGAGATGAGGATCCTGAAACAAAACGACAGAAATAGTAACATGTTTCTGTA 2293
QY 703 IlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyr 722
DB 2294 TTGTTTCTAGTCTCTGGAATATTTCTTTTATACATTTTCTCCAGGCTTCACATTT 2353
QY 723 GlyArgAlaGlyIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeu 742
DB 2354 GGCAAGCTGGGAGATCCTCACTAAGCGGCTTCGATACATGATGTTTTTCAGATCCATGCTG 2413
QY 743 TyrGluAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyLeuThrThrIle 762
DB 2414 AGACAGGATGTACGTGGTTGATGACCCCTAAACACCAACACCTGGAGCATTTGACAAACAGG 2473
QY 763 LeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThr 782
DB 2474 CTTGCCAATGATCGGCTCAAGTTAAAGGGCTATAGTTCCAGGCTTGTCTGTCATTACC 2533
QY 783 GlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMet 802
DB 2534 CAGAAATAGCAAACTCTGGACAGGCAATTTATATCTTATCTTATCTATGTTGGCAATTA 2593
QY 803 ThrPheLeuLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAla 822

Db 250 ATGGTGGTGGAACTTTGGCTGCCATCATCCATGGAGCTGGACTTCCTCTCATGATGCTG 309
Qy 21 ValLeuGlyGluMetSerAspAsnLeuLeuSer-----GlyCys 33
Db 310 GTGTTTGGACATGACGGATACCTTTGCAAAATGCGAGAAATTTAGGAGATTTAGGAGCT 369
Qy 34 LeuValGlnThrAsnThrTySer----- 41
Db 370 CTGTTGTTTAAACACACTAATAGCAGTAATATCACTGATACAGTGGCCGTCATGAATCTG 429
Qy 42 -----PhePheArgLeuThrLeuTyTyValGlyIleGlyValAlaLeuLeu 58
Db 430 GAGGAAGATATGACCATGTCCTATTATACAGTGGAAATGGTGTGGGTGCTGGTT 489
Qy 59 PheGlyTyIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIle 78
Db 490 GCTGCTTACATTCAGGTTTCATTGTTGGTGGCTGGACGCTGGAAGACAAATACACAAAT 549
Qy 79 ArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAsp 98
Db 550 AGAAAAACAGTTTTTCATGCTATTAATCGACAGGATAGCTGGTTTGTATGTCACGAT 609
Qy 99 IleGlyGluLeuAsnThrArgMetThr-----AspIleAspLysIleSerAspGlyIleGly 117
Db 610 GTTGGGAGCTTAAACCCCGCTTTACAGATGATGCTCTCCAAGATTATGAAGAAATGGT 669
Qy 118 AspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGly 137
Db 670 GACAAAATTTGAATGTTCTTTCACTCAATGGCAACATTTTTCAGTGGTTTATAGTAGGA 729
Qy 138 LeuValLysGlyTyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAla 157
Db 730 TTTACAGCTGGTGGAGACTAACCTTGTGATTTGGCCATCAGTCTGTTCTTGGACTG 789
Qy 158 SerAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyI 177
Db 790 TCAGCTGCAGTCTGGCAAGATACTGCTCTTCAATTTACTGATAAAGAACTCTTAGCTTAT 849
Qy 178 SerLysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPhe 197
Db 850 GCAAAAGCTGAGCAGTAGCTTGAAGAGGTCTTGGCAGCAATTAGAATGATTTGCAAT 909
Qy 198 ArgAlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyAlaTrp 217
Db 910 GGAGGACAAAAGAAAGAACTCGAA----- 933
Qy 218 PheTyPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyThrGlnAsn 237
Db 934 -----AGGTACAAACAAAAT 948
Qy 238 LeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGly 257
Db 949 TTAGAAGAGCTTAAAGAAATTTGGGATTAAGAAAGCTATTACAGCCAATATTCTATAGGT 1008
Qy 258 AlaValTyPhePheMetAsnGlyThrTyGlyLeuAlaPheTrpTyGlyThrSerLeu 277
Db 1009 GCTGCTTCCTCTTACTATCATGCTATTATGCTCTGGCCCTCTGGTATGGACACACCTG 1068
Qy 278 IleLeuAsnGlyGluProGlyTyThrIleGlyThrValLeuAlaValPhePheSerVal 297
Db 1069 GTCCTCTCAAAGGAA-----TATTCTATTGGCAAGTACTCACTGATTTCTTTCTGTA 1122
Qy 298 IleHisSerSerTyCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAla 317
Db 1123 TTAATTTGGGCTTTTAGTGTGGACAGGCATCTCCAAGCATTTGAAGCATTTGCAAAATGCA 1182
Qy 318 ArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPhe 337
Db 1183 AGAGGACAGCTTTTGAATCTTCAAGATATATTGAATAAGCCAAAGTATTGACAGCTAT 1242
Qy 338 SerThrAlaGlyTyLysProGluSerIleGlyThrValGluPheLysAsnValSer 357
Db 1243 TCGAAGAGTGGGCAACAAACCAAGATAATTAAGGGAAATTTGGAAATTCAGAAATGTTTCA 1302

Qy 358 PheAsnTyProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLys 377
Db 1303 TTCAGTTACCATCTCGAAAGAGTTAAGATCTTGAAGGGCTGAACTGAGGTGCGAG 1362
Qy 378 SerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGln 397
Db 1363 AGTGGGCAGACGCTGCTGGTTGGAAACAGCGGCTGTGGGAAGAGACAAACGCTCCAG 1422
Qy 398 LeuLeuGlnArgLeuTyAspProAspAspGlyPheIleMetValAspGluAsnAspIle 417
Db 1423 CTGATCGACAGGCTTTATGACCCACAGAGGCGATGTCAGTGTGTTGATGACAGGATATT 1482
Qy 418 ArgAlaLeuAsnValArgHisTyArgAspHisIleGlyValValSerGlnGluProVal 437
Db 1483 AGGACCATTAACGTAAGTTTCTACGGGAAATCATCGGTGTGTTGAGTCAGGAACCTGTA 1542
Qy 438 LeuPheGlyThrThrIleSerAsnAsnIleLysTyGlyArgAspAspValThrAspGlu 457
Db 1543 TTTGTTGCCACACGATAGCTGAAACATTCGCTATGTCGTGTCGTGAAGATGTCACCATGGAT 1602
Qy 458 GluMetGluArgAlaAlaArgGluAlaAsnAlaTyAspPheIleMetGluPheProAsn 477
Db 1603 GAGATTGAGAAAGCTGTCAAGGAAGCAATGCTTATGATTTATCATGAAATGCTCCTCAG 1662
Qy 478 LysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArg 497
Db 1663 AATTTTGACACCTGTTGGAGAGAGGGCCACGTCAGTGTGGGCAAGAGCAGAGG 1722
Qy 498 IleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThr 517
Db 1723 ATCGCATTTGCAGTGCCTGTTGTCGCAACCCCAAGATCTCTCTGTCGACGAGGCCACG 1782
Qy 518 SerAlaLeuAspSerGluSerLysSerAlaValGlnAlaLeuGluLysAlaSerLys 537
Db 1783 TCAGCTTGACACACAAAGTGAAGCAGTGTTTCAGTGTGCTCTGGATAGGCCAGAAAA 1842
Qy 538 GlyArgThrThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIle 557
Db 1843 GGTTCGGACACCATTTGTATAGTCTCATCTGTTGCTACGGTTCGTAATCCGACGTCATC 1902
Qy 558 ValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLys 577
Db 1903 GCTGGTTTCGATGATGAGTCAATTTGTGGAGAAAGGAAATCATGATGAGCTCATGAAGAG 1962
Qy 578 ArgGlyLeuTyThrSerLeuValMetSerGlnAspIle----- 590
Db 1963 AAGGCATTTTACTTTCAAACTTTGCACAAATGCACACAGCAGGAAATGAAATGATAGAA 2022
Qy 591 LysLysAlaAspGlu-----GlnMetGluSerMetThrTySerThrGluArgLys 607
Db 2023 AATGCAGCTGATCAATCCAAAGTGAATTTGATACCTTGGAAATGCTTCCACATGATTCA 2082
Qy 608 ThrAsnSerLeu-----ProLeuHisSerValLysSerIleLysSerAsp 622
Db 2083 GGATCCAGTCTAATAAGAAAGAAAGATCCACTCGTAGGAGTGTCCGTGGATCACAAGGCCAA 2142
Qy 623 PheIleAspLysAlaGluGluSerThrGlnSerLysGluLysSerLeuProGluValSer 642
Db 2143 -----GACAGAAAGCTTAGTACAAAGAGGCTCTGGATGAAAGATGATCTCCAGTTTCC 2196
Qy 643 LeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeu 662
Db 2197 TTTTGGAGGATTATGAAGCTAAATTTAACTGAGTGGCCTTATTGTTGTTGTTGTTT 2256
Qy 663 AlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIle 682
Db 2257 TGTGCCATTATAAATGGAGGTCTGCAACCCAGCATTTGTCAGTAATATTTTCAAAGATTATA 2316
Qy 683 ThrMetPhe---GlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTySer 701
Db 2317 GGGATTTTCAAGAAATGATGATGTCGGAACAAACACAGAAATAGTAGTAACTGTTTCA 2376

QY 702 MetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPhe 721
Db 2377 CTATTTCTTTCTAGCTCTCGAATTTGTTCTTTTATTACATTTTCTCCTCAGGGCTTCACA 2436
QY 722 TyrGlyArgAlaGlyLeuIleLeuThrMetAtqLeuArgHisLeuAlaPheLysAlaMet 741
Db 2437 TTTGGCAAGCTGGAGAGATCTCACAAAGCGGCTCGATACATGGTTTTCATCCATG 2496
QY 742 LeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThr 761
Db 2497 CTCAGACAGATGTGAGCTGGTTTGATGACCTCAAAACACCACTGGAGCATTGACTACC 2556
QY 762 IleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeu 781
Db 2557 AGGCTGCCAATATGCTGTCAAGTTAAAGGGCTATAGTTTCCAGGCTGTCTAATTT 2616
QY 782 ThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyrGlu 801
Db 2617 ACCCAATATACCAATCTTGGGACAGGAAATATATATCTTAATCTATGTTGGCAA 2676
QY 802 MetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThr 821
Db 2677 CTGACACTGTACTCTTAGCAATTTGTACCATCATTCATGCAATAGCAGGAGTTGTGAATG 2736
QY 822 AlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIle 841
Db 2737 AAAATGTTGTCTGGCAAGCACTGAAAGATAAGAAAGAACTAGAAAGTGTGGGAAGATC 2796
QY 842 AlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPhe 861
Db 2797 GCTACTGAAACATAGAAACTTTCCGAACCTGTTGTTCTTTGACTCAGGAGCAAGATTT 2856
QY 862 GluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGln 881
Db 2857 GAACATATGATGATCAGAGTTTGAGGTACCATACATACAGAAACTCTTTGAGGAAGCACAC 2916
QY 882 IleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGly 901
Db 2917 ATCTTTGGAATCAGCTTTCTTCCAGCAGGCAATGATGATTTTCTATGCTGATGT 2976
QY 902 PheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIle 921
Db 2977 TTTCCGTTTGGAGCCCTACTTGGTGACATAGTCTCATGAGCTTTCAGGATGTTCTGTGA 3036
QY 922 ValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaPro 941
Db 3037 GTATTTTACGCTGTGCTTTGTTGTCATGCGCGTGGGCAAGTCAGTTCATTTGCTCCT 3096
QY 942 GluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysPro 961
Db 3097 GACTATGCCAAAGCCAAAGTATCAGCAGGCCCATCATCATCATGATTCATTGAAAAACCCCT 3156
QY 962 AsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGlu 981
Db 3157 TTGATTTGACAGCTACAGCACAGAAGGCTTAAGCCGGAACATTTGGAAGAAATGTCA 3216
QY 982 PheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeu 1001
Db 3217 TTTAATGAAGTTGTTTCAACTATCCACCCGATCGGACATCCAGTGCTTCAGGGCTG 3276
QY 1002 SerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLys 1021
Db 3277 AGCTGGAAGTGAAGAGGCCAGACGCTGCTGCTGGGCGAGCATGCTGTGGGAAG 3336
QY 1022 SerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPhe 1041
Db 3337 AGCAGCGTGTCTCAGCTCTCTGGAGCGGTTCTATGACCCCTTGGCGGGAAGTGTGCTT 3396
QY 1042 AspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleVal 1061
Db 3397 GACGGCAAGAAATAAAGCAACTGAATGTTTCAGTGGCTCCGAGCAACCTGGGCGATGCTG 3456
QY 1062 ProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsn 1081

Db 3457 TCCAGGAGCCCATCTCTGTTGACTGCGACATAGTGAGAACATTGCCCTATGAGACAAC 3516
QY 1082 SerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSer 1101
Db 3517 AGCCGGTGTCTCACAGGAGAGATCGTGAGGCGAGCCCAAGGAGCCATATACAGCC 3576
QY 1102 PheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeu 1121
Db 3577 TTTCTGAGTCACTGCTTAATAATATATAGCACAGATAGGAGACAAAGAACTCAGCTC 3636
QY 1122 SerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuGlnLysProLysIle 1141
Db 3637 TCTGGTGGCCAGAAACACGCAATGCCATAGCTCGTGCCCTTGTAGACAGCCCTCATATT 3696
QY 1142 LeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHis 1161
Db 3697 TTGCTTTGGATGAGCCACATCAGCTCTGGATACAGAAAGTGAAAGGTGTTCAGAA 3756
QY 1162 AlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAla 1181
Db 3757 GCCCTGGACAAAGCCAGAGAGGCGGTACCTGATTTGCTCACCCTGTCCACC 3816
QY 1182 IleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGlnGlyThr 1201
Db 3817 ATCCGAATGCAACTTAATAGTGTGTTTTCAGATGGCAGAGTCAAGGAGCAGGCACA 3876
QY 1202 HisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
Db 3877 CATCAGCAGCTGCTGGCAGAAAGGCATCTATTTTCATGGTCAAGTGTCCAGGCT 3933
RESULT 10
AF257746 4927 bp mRNA linear ROD 09-MAY-2000
LOCUS Rattus norvegicus multidrug resistance protein 1a (Pgyl) mRNA,
DEFINITION complete cds.
ACCESSION AF257746
VERSION AF257746.1 GI:7739772
KEYWORDS
SOURCE Rattus norvegicus.
ORGANISM Rattus norvegicus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 4927)
AUTHORS Hooiveld,G.J.E.J., Wilms,J.W.J., Hagenbuch,B., Jansen,P.L.M.,
Meijer,D.K.F. and Muller,M.
TITLE Cloning and functional characterization of the rat multidrug
resistance protein Mrdr1
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 4927)
AUTHORS Hooiveld,G.J.E.J., Wilms,J.W.J., Hagenbuch,B., Jansen,P.L.M.,
Meijer,D.K.F. and Muller,M.
TITLE Direct Submission
JOURNAL Submitted (19-APR-2000) Div. Gastroenterology & Hepatology,
University Hospital Groningen, Hanzeplein 1, Groningen NL-9713 GZ,
The Netherlands
FEATURES
source Location/Qualifiers
1. .4927
/organism="Rattus norvegicus"
/strain="Wistar"
/db_xref="taxon:10116"
/sex="male"
/cell_type="hepatocyte"
/tissue_type="liver"
/dev_stage="adult"
1. .4927
/gene="Pgyl"
86. .3904
/gene="Pgyl"
/note="ATP-binding cassette transport protein; Abcb1a"
/codon_start=1
/product="multidrug resistance protein 1a"


```
/protein_id="AAF69007.1"
/db_xref="GI:7739773"
/translation="MELELDNGRADKNFSKMGKSKKKEKKKPAVSLVLTWFRVAGW
LDRFVLTGLTALIHGIALPLMLVFGDMTDSFANVGNRSMFYNATDIYAKLEDE
MTYAYYYGNGVGLIYAIQVSLWCLAAGRIKIRKQFPFHAINNQLGIFDHDV
GEUNLTDVDSKINGIDGKGMFFQAMATFFGGFIIGTRGKULVLIALSPVIG
LSAGIKWALISSTFDKQAYAKAGAVEILAAITRTVAFGGKKELERYNNLBEA
KRIGIKATITANSMGAFLIYASVALAFWYGTSLVIKVEYTGQVLVFFSVLIGA
YPSKQASPIEAFANARGAAVEFPIIDNKPISIDFSKSHKPDNIQGNLEFNTHFS
YPSKQVQILKGLNVLKVGSTVALVNSGCGKSTVQLLQRLYDPIEGEVSIDGDI
RTINRYLREIIGVQSEPVLPATTIAENIRYGRNVNTWDEIEKAVKANANAFINKL
PHKFDLTVGERAGLGGQQRQAIARALVRNPPIILLDEATSALDTESEAVVQAALD
KAREGRTTVIAHRLSTVRNADIAGDGVIVFQGNHDELMREKGYFKLVMTQTAG
NETELGNCEACESGKIDNDVMSKSGSLIRRRSTRKIRGPHDDGDSLSKEALDD
DVPASFWRTILKLNSTWVPVVFVFCAILINGLOPAFSIIIPSKVGVFTKNDTPRIO
QNSNLSFLFLILIGIISPIITFLOQFTFGKAGEILTKRLRYMFKMLRQDISFDD
PKMTGALTIVVMDAAQVKGATSLAVITQNIANGILISLIYSGQLTLLAL
VPIIATAGVVMKMSGLQKOKKELEGSKTAEIENPRTVLSVTRQSEKTVIYAO
SLQIPYRNALKKAVHFGITFTFTQAMMYFSYACFRFGAYLVARELNFENLVLSA
IVFGMAVGQVSFPDYAKAKVSASHIIRIIEKIPEIDSYSTEGMLKPNLGNVKFN
GUMFNPTRPNIPLVGLSLVKKGGOTLALVSGSGCGKSTVQLLRFYDPMAGTVFL
DGEIKQLNOMLRHLGIVSQEPILFDCSIENIAYCNSRVSVSHEEIVKAKEANI
HQPIDSPKYNTRNGDKGTQSGGQKRIARALVRPHILLDEALDTESEK
VVOEALDKAREGRTCTVIAHRLSTIQNADLIIVIQNGQVKEHGHQQLLAQKGIYFSM
VSQAGAKRS"
```

BASE COUNT 1474 a 1037 c 1190 g 1226 t
ORIGIN

Alignment Scores:

Pred. No.: 4,24e-227 Length: 4927
Score: 3532.00 Matches: 682
Percent Similarity: 74.14% Conservative: 244
Best Local Similarity: 54.60% Mismatches: 263
Query Match: 57.35% Indels: 60
DB: 10 Gaps: 10

US-09-873-409-5 (1-1222) x AF257746 (1-4927)

```
Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 233 ATGCTGCTGGGAACCTTGCGGGCCATTCATCCATGAATGTCCTCCACCTTATGATGCTG 292
Qy 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
Db 293 GTCCTTTGGAGACATGACAGATAGCTTTTGCAAAT-----GTAGGAACAACCGTAGTATG 346
Qy 41 SerPhePheArgLeuThr----- 46
Db 347 AGTTTCTACAAATGTCACACATATATGCCAAGCTGGAGGACGAAATGACCCAGTCAGCC 406
Qy 47 LeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeu 66
Db 407 TACTATTACACGGGCACTGGGCGGTGTGCTCATGCTTGCCTCATCTCCAGGTTTCACTT 466
Qy 67 TrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGlnPhePheHisSerVal 86
Db 467 TGGTGCCTGGCAGCTGGGAGACAAATACACAAATAGGACAGAGTTTTCATGCCATC 526
Qy 87 LeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGluLeuAsnThrArgMet 106
Db 527 ATGAATCAGGAGATAGGCTGTTTGACGTGTCATGACGTGGGGAGCTCAACACCCGGCTC 586
Qy 107 Thr---AspIleAspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGln 125
Db 587 ACAGATGACGCTCCAAATAATAGAGGAATGGTGACAAAATGGAAATGGAAATGTTCTTTCAG 646
Qy 126 AsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLysGlyTyrLysLeuThr 145
Db 647 GCAATGGCAACATTTTGGTGGTTTATATAGATTACTCTCGCGCTGGAGACTAACT 706
Qy 146 LeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAlaCysSerArgMet 165
Db 707 CTTTGATTTTGGCCATCAGCCCTGTTCTTGGACTGTGAGCCACGCTTGGACACAGAAAGGAA 1759
```

```
Qy 166 ValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGlu 185
Db 767 TTGCTCTTACTTACTGATTAAGGAACCTCCAGGCTTATGCAAAAGCTGGAGCAGTTGCTGAA 826
Qy 186 GluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGlnLysGluLeuGln 205
Db 827 GAAGTCTTACGACATCAGAACTGTGATTCCTTTGGAGGACAAAAGAGGAACCTTGA 886
Qy 206 ArgSerPheLeuLeuAsnIleThrArgTyrAlaTyrPheTyrPheProGlnTyrPheLeuLeu 225
Db 886 ----- 886
Qy 226 SerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 245
Db 887 -----AGGTACAATAACAATTTGGAAGAAGCTAAAGAGCTTGGG 925
Qy 246 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGly 265
Db 926 ATAAAGAAGACTATCACGGCCAAACATTTCCATGGGTGCAGCTTTCTGCTTATCTATGCA 985
Qy 266 ThrTyrGlyLeuAlaPheThrTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 285
Db 986 TCATATGCTCTGCGATTCGTGATGGACTTCTTGGTCATCTCAAAAGAA-----TAC 1039
Qy 286 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 305
Db 1040 ACTATTGGCAAGTGTCTACTGCTTTTTCGTATTAATGGAGCATTCAGTGTGGG 1099
Qy 306 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaPheHisIlePhe 325
Db 1100 CAGGCACTCTCCAAATATTGAAGCTTTCGCAATAGCTAGAGGAGCAGCTTATGAAGTCTTC 1159
Qy 326 GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 345
Db 1160 AGTATAATTAATAAAGCCAGTATAGACAGCTTCTCAAAAGAGTGGGCAACACCCGAC 1219
Qy 346 SerIleGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 365
Db 1220 AACATACAGGAATTTGGAATTCANAAATATTCCTTCACTTACCGTCTCGAAAGAC 1279
Qy 366 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 385
Db 1280 GTTCAGATCTTGAAGGCTCTCAACCTGAAGGTGAAGCGGCGGACGCTAGCTGCTGTT 1339
Qy 386 GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro 405
Db 1340 GGCAACAGTGGCTGTGGGAAAAGCACAACTGTCTCAGTGTCTGAGAGGCTCTACGACCCC 1399
Qy 406 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 425
Db 1400 ATAGAGGCGAGGTGATCTCGACGGACAGGACATCAGGACCATCAATGTGAGGTATCTG 1459
Qy 426 ArgAspHisIleGlyValValSerGlnGluProValIlePheGlyThrThrIleSerAsn 445
Db 1460 CGGGAATCATTTGGGTGGTGGTTCAGTTCAGGAACCCGCTGCTGTTGCCACCAATTCGGAA 1519
Qy 446 AsnIleLysTyrGlyArgAspValThrAspGluLeuMetGluArgAlaAlaArgGlu 465
Db 1520 AACATTGCTATCGCGAGAAAACGTCACCATGGATGAGATAGAGAAAGCTGTCACAGGAA 1579
Qy 466 AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 485
Db 1580 GCCAATGCTATGATTCATCATGAACCTGCCCCACAAATTTGCACCCCTGGTGGTGAG 1639
Qy 486 LysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal 505
Db 1640 AGAGGGCGCAGCTAGTGGGGGACAGAAACAGAGGATCGCCATTCGCCGGGCTCTGGCT 1699
Qy 506 ArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 525
Db 1700 CGCAACCCCAAGATCCTTTTGTGGATGAGGCCACGCTCAGCCTTGGACACAGAAAGGAA 1759
```


QY 526 SerAlaValGlnAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAla 545
DB 1760 GCGTGGTTCAGGCGCTCTGGTAAGGCTAGAGAGCGCGACCAATGTGATGACT 1819
QY 546 HisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeu 565
DB 1820 CACGCTTGCTACAGTTCCGAATCGCATGCTCATCTCGTGTGTTGATGGTGTGCAAT 1879
QY 566 AlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuVal 585
DB 1880 GTGAGCAAGAAATCATGATGAGTCATGAGAGAGAAAGAAATTTACTTCAAACITGTC 1939
QY 586 MetSerGln-----AspIleLysAlaAspGluGlnMetGluSer----- 599
DB 1940 ATGATCAGACAGCAGGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 1999
QY 600 -----MetThrTyrSerThrGluArgLysThrAsnSerLeu-----ProLeu 613
DB 2000 GGAATTGATGATGTGGACATGCTCTCAAAAGATTTCAGGATCCAGTCTAATAAGAAAGA 2059
QY 614 HisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThr---Gln 632
DB 2060 TCAACTCGCAAAAGCATCCGTGGCGGCACATGATCAAGACCGGGAACTTAGCACCAAGAG 2119
QY 633 SerLysGluLeuSerLeuProGluValSerLeuLysIleLeuLysLeuAsnLysPro 652
DB 2120 GCTGTGATGACGACGTACCTCCAGCTTCTCTTTGGCGGATCCCTGGAAGTTGAAATCAACT 2179
QY 653 GluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPro 672
DB 2180 GAATGGCTTATTGTTGGTGTGTGATTGTTGGCCATAATAATGAGGCTTGCACCA 2239
QY 673 ValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThr 692
DB 2240 GCATCTCCATAATATTTCAAGAGTTGTAGGGGTTTTACAAAATGACACCCCTGAA 2299
QY 693 Leu---LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCys 711
DB 2300 ATCAGCGGCAGAACACAACTTCAAGAGTTGTAGGGGTTTTACAAAATGACACCCCTGAA 2359
QY 712 PheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMet 731
DB 2360 TTCAATACGTTTTCCTTCAGGGCTTCACATTTGGCAAGCTGGAGAGATCCACCAAG 2419
QY 732 ArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGlu 751
DB 2420 CGACTCGATACATGCTTCAATCCATGCTGAGACAGACATAGCTGGTTGATGAC 2479
QY 752 LysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGln 771
DB 2480 CCTAAAACACCAAGGCGCTGACCAAGGCTTCCCAATGACGCTGCTCAAGTGAA 2539
QY 772 GlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSer 791
DB 2540 GGGGCTACAGGGTCTAGGCTGTGTTATTACCCAGAACATAGCAATCTTGGGACAGGC 2599
QY 792 ValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaPro 811
DB 2600 ATCATCATATCCCTGATCTAGGCTGGCAATTGACATTTTACTCTAGCAATGTTTCCC 2659
QY 812 ValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAsp 831
DB 2660 ATCATTCCTATAGCAGGAGTCTGGGAAGATCGCTACAGAACAAATGAAACTTTCCGACT 2719
QY 832 LysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThr 851
DB 2720 AAGAGGAACCTAGAGGTTCTGGGAAGATCGCTACAGAACAAATGAAACTTTCCGACT 2779
QY 852 IleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThr 871
DB 2780 GTGCTCTTTGATCGGGAGCAGAGATTGAAACTATGATATGCCAGAGCTTGCAGATA 2839
QY 872 GlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCystyrAlaPheSerHis 891

DB 2840 CCATACAGAAATGCTTTGAAGAAAGCGCGCTTTTGGGATCATCTTCTCTTCACCCAG 2899
QY 892 AlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAla 911
DB 2900 GCCATGATGATTTCTCTATGCTGCTGTTTCGGGTTTGGTGCCTACTTGGTGGCAGCA 2959
QY 912 GlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMet 931
DB 2960 GAACATCATGACATTTGAAATATGTTCTGTAGTATTCTCAGCTATTGCTTGTGGTCATG 3019
QY 932 AlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAla 951
DB 3020 GCAGTGGCGAGCTCAGTTTCATTCGCTCTGACTACGCCGAAAGCCAAAGTCTCGGCATCC 3079
QY 952 HisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLys 971
DB 3080 CACATCATCAGGATCATTCAGAAATCCCTGAGATTGACAGCTACAGCAGGAGGCTTG 3139
QY 972 LysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCys 991
DB 3140 AGCCCTAATATGTTGAAGGAAATGTAAATTTAATGGAGTCATGTTCAACTATCCACC 3199
QY 992 ArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrVal 1011
DB 3200 CGACCCCAACATCCAGTGTCTCAGGGGCTGAGCCTAGAGGTGAAGAAAGCGCAGCGCTG 3259
QY 1012 AlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeu 1031
DB 3260 GCCTCTGCTGGCAGCAGTGGCTGCGGAGAGAGTACAGTGGTCCAGCTGCTTGAGCGCTTC 3319
QY 1032 TyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnVal 1051
DB 3320 TATGACCCCATCCGCGGAAACAGTGTCTTAGATGGCAAGAAATAAAGCAACTCAATGTC 3379
QY 1052 GlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSer 1071
DB 3380 CAGTGGCTCGCGCCACCTCGGCATTTGTGCCAGAGCCCATCTCTGTTGATCGAGC 3439
QY 1072 IleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLys 1091
DB 3440 ATGCCGAGAACATTCCTACGGAGAACACAGCCGTGCTGTCTCATGAGGAGATCGTG 3499
QY 1092 GluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsn 1111
DB 3500 AAGCGACCAAGAGGCGCAACATCCACAGTTTCATCGACTCACTGCTGAGAAATACAA 3559
QY 1112 ThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIle 1131
DB 3560 ACCAGAGTGGGAGACAAAGGAGCTCAGCTGTCCGGCGGCGAGAGCGCATCGGCATC 3619
QY 1132 AlaArgAlaLeuLeuGlnLysProLysIleLeuLeuAspGluAlaThrSerAlaLeu 1151
DB 3620 GCGCGGCCCTGCTACAGACGCTCATCTTACTTCTGTGATGAAGCAGCATCAGCTCTG 3679
QY 1152 AspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThr 1171
DB 3680 GATACGAGAGTGAAAGGTCTGTCAGGAAGCGCTGCAAAAGCCAGGAGGCCGCCACC 3739
QY 1172 CysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValLeu 1191
DB 3740 TGCATTTGATCGCGCAGCGCTGTCACCATCCAGAACGCGACACTTGATCGTGGTGATT 3799
QY 1192 HisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuArgAsnArgAspIle 1211
DB 3800 CAGACGGCAGGCTCAAGAGCAGCGCACCCACAGCAGCTGCTGCCCAAGAAAGCATC 3859
QY 1212 TyrPheLysLeuValAsnAlaGlnSer 1220
DB 3860 TATTTCTCGATGCTCAGTGTGCGGCT 3886
RESULT 11
AX322787

LOCUS AX322787 3860 bp DNA linear PAT 07-JAN-2002
DEFINITION Sequence 1 from Patent WO0192877.
ACCESSION AX322787
VERSION AX322787.1 GI:18093766
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Sorrentino, B. and Schuetz, J.
TITLE Method of identifying and/or isolating stem cells
JOURNAL Patent: WO 0192877-A 1 06-DEC-2001;
ST. JUDE CHILDREN'S RESEARCH HOSPITAL (US)
FEATURES
source 1..3860
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 1135 a 746 c 957 g 1022 t
ORIGIN
Alignment Scores:
Pred. No.: 3.9e-227 Length: 3860
Score: 3530.50 Matches: 682
Percent Similarity: 74.00% Conservative: 246
Best Local Similarity: 54.39% Mismatches: 263
Query Match: 57.32% Indels: 63
DB: 6 Gaps: 10
US-09-873-409-5 (1-1222) x AX322787 (1-3860)
Qy 1 MetileLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 151 ATGGTGGGAACTTTGGCTGCCATCATCCATGGGCTGGACTTCCTCTCATGATCGT 210
Qy 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
Db 211 GTGTTTGGAGAAATGACAGATATCTTTGCAAAATGACAGAAATTTAGAGATCTGATGCA 270
Qy 32 GlyCysLeuValGlnThr-----AsnThrTySerPhePhe----- 43
Db 271 AACATCACTAATAGAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAC 330
Qy 44 -----ArgLeuThrLeuTyTyValGlyIleGlyValAlaLeuIlePheGlyTy 61
Db 331 ATGACAGATATGCTATTATTACAGTGGATTTGGTCTGGGGTCTGGTTCCTGCTTAC 390
Qy 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db 391 ATTCAGGTTTCATTTGGTGGCTGGCAGCTGGAGACAAATACACAAAATTTAGAAAACAG 450
Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db 451 TTTTTCATGCTATATATGCGACAGAGATAGCTGGTTGATGTCACGATTTGGGGAG 510
Qy 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db 511 CTTAAACCCGACTTACAGATGATGCTCTAAGATTAAATGAAGTATTTGGTGCACAAAT 570
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db 571 GGAATGTTCTTCAAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGGATTTACACGT 630
Qy 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db 631 GGTTGGAGCTAAACCTGTGATTTTGGCCATCAGTCTCTGTTCTGGACATGTCAGCTGCT 690
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTySerLysAla 180
Db 691 GTCTGGCAGAGATACATCTTCTTACTGATAAAGACTCTTAGGATGATGCAAAAGCT 750
Qy 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db 1744 ACCATTTGTATAGTCTCATCGTTTGTCTACAGTTCTGTAAGTCTGACGTCATCGTGGTTTC 1803

Db 751 GGAGCAGTAGCTAGCGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTCATTTGGAGACAA 810
Qy 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
Db 811 AAGAAAGAACTTGAA-----AGGTACAAACAAAATTTAGAAAGAA 825
Qy 221 ProGlnTrpLeuLeuSerCysValLeu**PheValArgTyrThrGlnAsnLeuLysAsp 240
Db 826 -----AGGTACAAACAAAATTTAGAAAGAA 849
Qy 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
Db 850 GCTAAAAGAAATTTGGGATAAAGAAAGCTATTACAGCAATATTTCTATAGGTGCTGCTTC 909
Qy 261 PhePheMetAsnGlyThrTyGlyLeuAlaPheTyTyGlyThrSerLeuIleLeuAsn 280
Db 910 CTGCTGATCTATGCACTCTTATGCTCTGGCTTCTGTTATGGGACCACTTGGTCTCTCA 969
Qy 281 GlyGluProGlyTyThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
Db 970 GGGGAA-----TATTCTATTGGCAAGTACTCACTGTATTCTTTCTGTATTAAATGGG 1023
Qy 301 SerTyTyCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaAsgGlyAla 320
Db 1024 GCTTTTAGTGTGGACAGGCATCTCCAAGCATTTGCAAAATGCAAGAGAGCA 1083
Qy 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
Db 1084 GCTTATGAATCTTCAAGATATTGATATAGCCAAAGTATTGACAGCTATTGCAAGAGT 1143
Qy 341 GlyTyTyProGluSerIleGlyThrValGluPheLysAsnValSerPheAsnTy 360
Db 1144 GGGCACAAACCATATAATTAAAGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 1203
Qy 361 ProSerArgProSerIleLysLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
Db 1204 CCATCTCGAAAAGAAAGTAAAGATCTTGAAGGGCTGAACCTGAAGGTGACAGATGGGCA 1263
Qy 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
Db 1264 ACCTGGCCCTGTTGGAAACAGTGGCTGGGAAGACACACAGTCCAGCTGATGTCAG 1323
Qy 401 ArgLeuTyAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
Db 1324 AGGCTCTATGACCCACAGAGGGGATGGTCAGTGTGATGGACAGGATATTAGGACCATA 1383
Qy 421 AsnValArgHisTyArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
Db 1384 AATGTAAGSTTTCTACGGGAAATCATTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1443
Qy 441 ThrThrIleSerAsnAsnIleLysTyTyGlyArgAspAspValThrAspGluMetGlu 460
Db 1444 ACCACATAGCTGAAAACATTCGCTATGGCCGTAAGATGTCACCATGGATGATGATGAG 1503
Qy 461 ArgAlaAlaArgGluAlaAsnAlaTyAspPheIleMetGluPheProAsnLysPheAsn 480
Db 1504 AAAGCTGTCAAGGAGCAATGCTATGCTTATCATGAAGTGGCTCATATAAATTGAC 1563
Qy 481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
Db 1564 ACCCTGGTGGAGAGAGGGGCCAGTTGAGTGGTGGGCGAGAGAGAGATCGCCATT 1623
Qy 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
Db 1624 GCACGTGCCCTGGTTGCAACCCCAAGATCTCTCTGCTGGATGGCCAGCTCAGCCTTG 1683
Qy 521 AspSerGluSerLysSerAlaValGlnAlaLeuGluLysAlaSerLysGlyArgThr 540
Db 1684 GACACAGAAAGCAAGCAGTGGTTTCAAGTGGCTCTGGATAAGGCCAGAAAAGGTCCGACC 1743
Qy 541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560
Db 1744 ACCATTTGTATAGTCTCATCGTTTGTCTACAGTTCTGTAAGTCTGACGTCATCGTGGTTTC 1803

QY 561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 1804 GATGATGGAGTCATGTGGAGAAAGGAAATCATGAACTCATGAAAGAGAAAGCAAT 1863
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
QY 581 TyrTyrSerLeuValMetSerGln-----AspIleLysLysAlaAsp 594
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 1864 TACTTCAAACTTGTCAATGATGCTTGGAAATGAAAGTGAAGTGAAGATGCAGCT 1923
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
QY 595 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis 614
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 1924 GATGAATCCAAAGTGAATGATGCTTGGAAATGCTCTCAATGATTCAGATCCAGT 1983
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
QY 615 SerVal-----LysSerIleLys---SerAspPheIleAspLysAla 627
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 1984 CTAATAAGAAAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAG 2043
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
QY 628 GluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeu 647
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 2044 CTTAGTAGTCAAAAGAGGCTCTGGATGAAGATATATACCTCCAGTTTCTTTGGAGGATATG 2103
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
QY 648 LysLeuAsnLysProGluTyrPheValValLeuGlyThrLeuAlaSerValLeuAsn 667
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 2104 AAGCTAAATTAACCTGAATGCCCTTAATTTGTTGGTGATTTTGTGCCATTAAT 2163
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
QY 668 GlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn 687
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 2164 GGAGGCTGCAACACGATTTGCAATAATATTTTCAAGATATAGGGTTTTCACAAGA 2223
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
QY 688 ---AsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIle 706
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 2224 ATTGATGATCCTGAAACAAACACACAGATAGTAACCTTGTTTTCACTATTGTTCTAGCC 2283
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
QY 707 LeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGly 726
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 2284 CTTGGAAATTAATTTTATTATATTTTCTTCAAGTTTTCATTTGGCAAGCTGGA 2343
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
QY 727 GluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIle 746
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 2344 GAGATCCTCAACAGCGCTCCGATACATGTTTCCGATCCATGCTCAGACAGGATGTG 2403
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
QY 747 AlaTyrPheAspGluLysGluAsnSerThrGlyClyLeuThrThrIleLeuAlaIleAsp 766
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 2404 AGTTGGTTTTCATGACCCCTAAACAAACACCATCTGAGCATTTGATCCAGGCTCGCCAAATGAT 2463
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
QY 767 IleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThr 786
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 2464 GCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTTCTGCTAATATCCCAAGATATAGCA 2523
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
QY 787 AsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyrGluMetThrPheLeuIle 806
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 2524 AATCTTGGGACAGGATTAATATATCTTCTCATCTATGTTGGCACTTAACACTGTACTC 2583
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
QY 807 LeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGly 826
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 2584 TTAGCAATTTGATCCCATCATTTGCAATAGCAGGAGTTGTTGAAATGAAATGTTGCTCGGA 2643
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
QY 827 PheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeu 846
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 2644 CAAGCACTGAAGATGAAGAAAGCACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATA 2703
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
QY 847 GluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGlu 866
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 2704 GAAACTTCGAACCGTTGTTTCTTGACTCAGAGGAGGAAGTTTGAACATATGATGCT 2763
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
QY 867 GluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCys 886
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 2764 CAGAGTTTGCAGGTACCACAGAAACTCTTTTGAGGAAAGACACACATCTTTGGAATTTACA 2823
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
QY 887 TyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAla 906
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 2824 TTTTCTTCAACCCAGGCAATGATGATTTTCTCTATGCTGGATGTTTCCGGTTTGGAGCC 2883
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
QY 907 TyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIle 926
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 2884 TACTTGGTGGCACAATAAACTCATGAGCTTTTGGAGTCTTCTGTAGTATTTTTCAGCTGT 2943
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
QY 927 AlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAla 946
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 2944 GTCTTTGGTGGCATTGCGCTGGGCAAGTCAAGTCACTTATTTGCTCTGACTATGCCAAGCC 3003
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
QY 947 LysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArg 966
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 3004 AAAATATCAGCAGCCACATCATCATGATCATTTGAAAAAACCCCTTTGATTGACAGCTAC 3063
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
QY 967 SerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSer 986
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 3064 AGCAGCAAGAGGCTTAATGCCGAACATTTGGAAGGAAATGTCACTATTTGGTGAAGTTGTA 3123
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
QY 987 PhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGlu 1006
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 3124 TTCAACTATCCACCCGACCGGACATCCAGTGTCTTCAGGACTGAGCCTGGAGGTGAAG 3183
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
QY 1007 ArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGln 1026
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 3184 AAGGCCACAGCTGGCTCTGTGGGSCACAGTGGCTGTGGGAAGACACAGTGGTCCAG 3243
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
QY 1027 LeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAla 1046
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 3244 CTCCTGGAGCGTCTTACGACCCCTTGGCAGGGAAGTGTCTGCTTGCATGGCAAGAAATA 3303
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
QY 1047 LysGluLeuAsnValGlnTyrPheArgSerGlnIleAlaIleValProGlnGluProVal 1066
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 3304 AAGCGACTGAATGTTTCAAGTGGCTCCGAGCACACCTGGGATCGTGTCCAGGAGCCATC 3363
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
QY 1067 LeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValPro 1086
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 3364 CTGTTGACTGCAGCATTTCTGAGAACATTTGCTATGGAGACACACAGCCGGTGTGTC 3423
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
QY 1087 LeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeu 1106
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 3424 CAGGAAGAGATCGTGAAGGCGAGCAAGAGGAGGCAACATATACATGCTTTCATCGAGTCACTG 3483
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
QY 1107 ProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLys 1126
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 3484 CCTATAAATATAGCACTAAAGTAGGAGACAAAGGAACCTCAGCTCTCTGTGTGCCAGAAA 3543
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
QY 1127 GlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGlu 1146
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 3544 CAACGATTTGCCATAGCTGTGCCCTTGTAGACACCTCATATTTGCTTTGGATGAA 3603
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
QY 1147 AlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAla 1166
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 3604 GCCACGTCAAGCTCTGATACAGAAAGTGAAGGTTGTCCAAGAGCCCTGGACAAAGCC 3663
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
QY 1167 ArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAsp 1186
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 3664 AGAAGAGGCGCAGCTGTCATTTGATTGCTCCAGCTGTCCACCATCCAGATGCGAGC 3723
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
QY 1187 LeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeu 1206
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 3724 TTAATAGTGGTGTTCAGATGGCAGATCAAGAGCATGGCACGCATCAGCAGCTGCTG 3783
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
QY 1207 ArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 3784 GCACAGAAAGGCATCTATTTTCAATGGTTCAGTGTCCAGGCT 3825
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
RESULT 12
AX105078
LOCUS AX105078 4279 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 22 from Patent WO0123540.
ACCESSION AX105078
VERSION AX105078.1 GI:13921228
KEYWORDS dog.
SOURCE Canis familiaris
ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.	
REFERENCE	1 (bases 1 to 4279)
AUTHORS	Stocker,P.J., Steimel-Crespi,D.T., Crespi,C.L., Reif,T.C. and Patten,C.J.
TITLE	P-glycoproteins and uses thereof
JOURNAL	Patent: WO 0123540-A 22 05-APR-2001;
FEATURES	GENTEST CORPORATION (US)
source	Location/Qualifiers 1..4279
CDS	/organism="Canis familiaris" /db_xref="taxon:9615" 17..3862 /note="unnamed protein product"
	/codon_start=1 /protein_id="CAC37721.1" /db_xref="GI:13921229"
	/translation="MDPEGRKSGAEKNFMWRKKSKNKKPKPTVSTFAMFRYSN WLDRLMLVGTMAAI THGAALPLMLVFNGMTDSFANAGISRNKTTPVI INESITNN QHFNHLEEMTYTAYYSIGGAGLVAAIYQSFMC LAAGROIILRIKQFPFAIMRO EICWFVDHVGEINLTIDVDSKINSEGIDKI GFWFFQS IATPFTGFIVTRGWOKL VLIAISPGLGSRAIIWAKLITANTISIGAFLII YASVALAFWYCTSLVSESYIGGV ERYNKLEAGKIGIKKAITANTISIGAFLII YASVALAFWYCTSLVSESYIGGV TVFSVLIEGASIQASPSTEANARGAAYEIFKIIDNKPSIDSYSKSHKPDNTKG LMFKNVHSPSRKEKILUGLNLIKVSOGQTVALVNSCGCKSTTVQLMORLYDPTD GMVICDODRTITNVHLREITGVSPQPVLFATTI AENIRYGRENVTMDEIEKAVKE ANAYDFMKLPNFEDTLVBGRQA LSCGQKQRTA IARALVRNP KILLILDATTSALDTE SEAVQMTRGNELELENATESKSDALEMPKDSSLIKRRTSRTHAPQQQR FKLVTFQTNELKLNTEVPVSWFRLKLNTEVPVFI CAIINGLGQPAFSIIFSRIIGI KLTGEDNETRONSNPFSLVLGLIISPIFTFFGQTFPGKAGEIILT KRURYVMVFRSM TRDEDPEKTRONSNPFSLVLGLIISPIFTFFGQTFPGKAGEIILT KRURYVMVFRSM LRQDVSWPDFDKNTTGALTRLANDRAOVKAIGSR LAVITONIANLGTCI IISLIVG WQUTLLLAIVPIIATAGVTEMKM LSQALKDKKELEGACKIATEALEINTVSVLSIR EQKFVEPIAQSPVYNLUKKAHIFGVFSITQAMMYI CCFRFGALVAINEFNMVN FOVLVLVAIVFGAMAVGVSPFADYAKAKVSAAHVIMI IEKSPLIDSYSHGLKRP NTLEGNVTVEFNVPYTRPDI PVLQGLEVKKGQTTALVGSGCCKSTVVQLLERF YDLAAGSLVDHGIEIKHLNWMLRHGLGV SQBPILFDCSIAENIAYGDNRSVVHDEE IMOAAXEANTHTIETLPKYNTRVGDKQTGSGQKORTAIRALVRQPHILLDEEA TSALDTSEKVQOEALDKAREGRTCIVIAHRLSTIQNALDIVVFPQNGKVKHEHQTHQL LAQGYFSWSVQAGNR"
BASE COUNT	1295 a 833 c 1008 g 1143 t
ORIGIN	
Alignment Scores:	
Pred. No.:	4, 45e-227 Length: 4279
Score:	3530.50 Matches: 682
Percent Similarity:	73.13% Conservative: 238
Best Local Similarity:	54.21% Mismatches: 269
Query Match:	Indels: 69
DB:	Gaps: 6

US-09-873-409-5 (1-1222) x AX105078 (1-4279)

Qy	1	MetIleLeuGlyIleLeuIleSerLeuValAsnGlyAlaCysLeuProIleMetProLeu	20
Db	167	ATGTTGTTGGGCAATGGCTGCATCATCCATGGAGCTGCATCCCTCTCATGATGGT	226
Qy	21	ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu	34
Db	227	GTTTTGGAAACATGACAGATAGCTTTGCCAATGCGAGGAATTCAGAAACAAACTTTT	286
Qy	35	-----ValGlnThrAsnThrTy-SerPheArg-----	44
Db	287	CCAGTTATAATTAATGAAAGTATTATCGAACAAATACACACATTTCATCAACCATCTGGAG	346
Qy	45	-----LeuThrLeuTy-rValGlyIleGlyValAlaIleLeuPhe	59
Db	347	GAGAAATGACACCGTATGCTATTATTACATGGGATCGTGCTCGGTGCTGGTGGCT	406
Qy	60	GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg	79
Db	407	GCATTACATCCAGGTTTCATTCTGGTGGCTGGCAGCAGGAAGACAGATATCTCAAAATTAG	466

Qy	80	LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyThrPheAspSerCysAspIle	99
Db	467	AAACAATTTTTTCATGCTATCTGCACGACGAGGATGGCTGGTTTCACGTGCATAGACGTT	526
Qy	100	GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp	118
Db	527	GGGGAGCTTAACACCCGGCTCACAGACGATGCTCCAAAATCAATCAAGGAATTCGCGAC	586
Qy	119	LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu	138
Db	587	AAAATGGAAATGCTTTTCAATCAATAGCAACATTTTTCACCGCTTTTATAGTGGGGTTT	646
Qy	139	ValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer	158
Db	647	ACACGTGGTTTGGGAAGCTAACCCCTTGATTTTGGCCATCAGCCCTGCTCTGGACATTCA	706
Qy	159	AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTrpSer	178
Db	707	GCGGCCATCTGGCCAAAGATACTATCTTCAATTACTGATTAAGAACTCTTGGCTATGCA	766
Qy	179	LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg	198
Db	767	AAAGCTGGACAGTACTGTAAGAAGTCTTAGCAGCAATCAGAACTGTGATTGCTTTGGA	826
Qy	199	AlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPhe	218
Db	827	GGACAAAGAAGAAGAACTTGAA-----	847
Qy	219	TyrPheProGlnTrpLeuLeuSerCysValLeu**PheValArgTyrThrGlnAsnLeu	238
Db	848	-----AGGTACAAACAAATAATTTA	865
Qy	239	LysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyVala	258
Db	866	GAAGAAGCTAAAGGAATGGGATPAAGAAAGCTATCAGGCCAACATTTCTATTGGTGCC	925
Qy	259	ValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeulle	278
Db	926	GCTTCTTATTGATCTATGCATCATATGCTCTGGCTTCTGGTATGGGACCTCCTTGGTC	985
Qy	279	LeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValle	298
Db	986	CTCTCCAGTGAA-----TATTCTATTGGCAAGTACTCACTGCTCTCTTCTTGATTATA	103
Qy	299	HisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArg	318
Db	1040	ATTGGGCGCTTTTAGTATTGACAGGCAATCCCAAGCATTTGAACATTTGCACGCAAGA	109
Qy	319	GlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSer	338
Db	1100	GGAGCAGCTTATGAATCTTTCAAGATAATTTGACAATAAACAAGCATTTGACAGCTATTGC	115
Qy	339	ThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPhe	358
Db	1160	AAGAGTGGACATAAACACAGATAATATTAAAGGGAATTTTGGAAATTCAAAATGTGTCACTTC	121
Qy	359	AsnTyrProSerArgProSerIleIleValLeuLysGlyLeuAsnLeuArgIleLysSer	378
Db	1220	AGTTACCCCTTCTGAAAGAAGTTTAGATCTTAAGGGTCTCAACCTGAAGGTTTCAGAGT	127
Qy	379	GlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeu	398
Db	1280	GGGCAGACAGTGGCGCTGGTTGGGACAGTGGCTGGCGGAAGAGCAGCACCGTGCAGCTG	133
Qy	399	LeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArg	418
Db	1340	ATGCAGAGGCTCTATGACCCACACAGATGGCATGGTGTGTATTGATGGACAGGACATTAGG	139
Qy	419	AlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeu	438
Db	1400	ACCAATAATGTAAAGGCATCTTCGGGAATAATTACTGGTGTGTAGTCAGGAGCTGTGTTG	145
Qy	439	PheGlyThrThrIleSerAsnAsnIleLvsTyrGlyArgAspAspValThrAspGluGlu	458

Db 1460 TTTGCCACCATGATGCTGAACATTTGCGTATGGCGCGAATAATGTCACCATGGATGAG 1519
QY MetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLys 478
Db 1520 ATTGAGAAAGCTGTTAAGGAGGCAATGCTATGATTTATCATGAACATACCTAATAA 1579
QY PheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgile 498
Db 1580 TTTGACACTCTGGTTGGAGAGAGGGGCCAGCTGAGTGGTGAGCAGAAACAGAGAATC 1639
QY AlaIleAlaArgAlaLeuValArgAsnProLysIleIleuIleLeuAspGluAlaThrSer 518
Db 1640 GCCATTGCTCGGGCCCTGGTTCCCAACCCCAAGATTTCTTGCTGGATGAGGCAACGTCA 1699
QY AlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGly 538
Db 1700 GCTCTGACACTGAAAGTGAAGCAGTGGTTGAGTGGCCCTGGATAAGGCCAGAAAGGC 1759
QY ArgThrThrIleValValAlaAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleVal 558
Db 1760 CGGACTACCATTTGATAGTCATCGTTTGTCTACAGTTCCGTAATGCCGATGTCATTGCT 1819
QY ThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArg 578
Db 1820 GGTTTTGATGATGGAGTCATTTGGAGAAAGGAAATCATGATCAACTCATGAAAGAGAG 1879
QY GlyLeuTyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGlu 598
Db 1880 GGCATTACTTCAAACTTGTCACAAATGCAG---ACAAGAGAAATGAAATGATGATTAGAA 1936
QY SerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSer 618
Db 1937 AATGCCACTCGTGAATCCAAAGTGAAGTGAATGCTTGGAAATG---TCTCCAAAGAT 1993
QY IleLysSerAspPheIle----- 624
Db 1994 TCAGGGTCCAGTTTAATAAAAGAGATCAACTCGCAGGAGTATACATGCACCACCAAGGC 2053
QY 625 ---AspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeu 643
Db 2054 CAAGACAGAAAGCTTGGTCAAAAGAGGACTTGATGAGAAATGTACTCCAGTTTCTTTC 2113
QY LeuLysIleLeuLysLeuAsnLysProGluTrpPheValValLeuGlyThrLeuAla 663
Db 2114 TGGAGGATTCTGAAGCTGAACCTCAACTGAATGCTTATTTTGTGGTTGTATATTTGT 2173
QY SerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThr 683
Db 2174 GCTATTATAACGGAGGCTTGCACACCAAGCATTTTCAATAATATTTTCAAGGATTATAGG 2233
QY MetPhe---GlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMet 702
Db 2234 ATCTTTACCGAGATGAGGATCTCGAATAAAGACGACAGAAATAGTAACATGTTTCTGTA 2293
QY IlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyr 722
Db 2294 TTGTTTCTAGTCTCTGGAATTAATTTCTTTATATACATTTTCTCCAGGCTTCACATTT 2353
QY GlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeu 742
Db 2354 GGCAAGCTGGGAGATCCTCACTTAAGCGGCTTCGATACATGTTTTCAGATCCATGCTG 2413
QY TyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrIle 762
Db 2414 AGCAGAGATGTCAGCTGGTTGATGACCTTAAACACCACTGGAGCATTTGACAAACAGG 2473
QY LeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgileGlyValLeuThr 782
Db 2474 CTTGGCCNATGATCGCGCTCAAGTTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTCATFACC 2533
QY GlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMet 802

Db 2534 CAGATATATAGCAAAATCTTGGACAGGCATATATATATCTTAATCTATGTTGGCAATTA 2593
QY ThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAla 822
Db 2594 ACATTTTACTCTTAGCAATTTGACCATCATTTGCAATGCAAGAGTGTGTTGAATGAAA 2653
QY AlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAla 842
Db 2654 ATGTTGCTCTGGCAAGACACTGAAAGATAAGAAAGAGCTAGAGGAGCTGGGAAGATTGCT 2713
QY ThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGlu 862
Db 2714 ACAGAACCATCGAAACATTCGAACTGTGTTCTTTGACTCGGAGCAGAGATTGAA 2773
QY GlnMetTyrGluGluMetLeuGlnHisArgAsnThrSerLysLysAlaGlnIle 882
Db 2774 TACATGTATGCACAGAGTTTGAAGTACCATACAGAAACTCTTTGAGGAAAGCACATC 2833
QY IleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPhe 902
Db 2834 TTCGGGGTCTCATTTTCTATCACCCAGGCAATGATATATTTTCTATGCTGGCTGTTTC 2893
QY ArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleVal 922
Db 2894 CGGTTTGGTCCCTACTTGTGGCAATGATTCATGAACTTTTCAGGATGTTCTTTTGGTA 2953
QY PheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGlu 942
Db 2954 TTTCTAGCTATTGCTCTTGTGTCATGGCAGTGGGCGAGTCACTGTTCAITTTGCTCTGAC 3013
QY TyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsn 962
Db 3014 TATGCCAAAGCCAAAGTATCAGCAGGCCACGTCATCATCATCATTTGAAAAAGCCCTCTG 3073
QY IleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluLysLeuLeuGluPhe 982
Db 3074 ATTGACAGCTACAGCCCTCAGCGCCTCAAGCCCAATACGTTGGAAGAAATGTGACATTT 3133
QY ArgGluValSerPheTyrProCysArgProAspValPheIleLeuArgGlyLeuSer 1002
Db 3134 AATGAGGTGCTGTTCAACTATCCCACTCGACCAGACATCCCGTGTCTCCAGGGGCTGAGC 3193
QY LeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSer 1022
Db 3194 CTGAGGTGAAGAAGGCCAGACGCTGGCCCTCGTAGGTAGCAGTGGCTGTGGGAAGAGC 3253
QY ThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAsp 1042
Db 3254 ACAGTTGTTGAGTCTCTAGAGCGCTTCTATGACCCCTTGGCTGGTTCAGTGTAAATTGAT 3313
QY GlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValPro 1062
Db 3314 GGCAAGAGATAAAGCACCTGAATGTCCAGTGGCTCCGAGCACACCTGGGCATCGTGTCT 3373
QY GlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSer 1082
Db 3374 CAGGAGCCCATCTCTGTTGACTGACGCAATTCGCGAACAATTCCTATGCGAGACAACAGC 3433
QY ArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPhe 1102
Db 3434 CGGGTCTGATTCATCAAGAGATTTATGAGGAGCCCAAGAGGCCCAATACACCACTTC 3493
QY IleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSer 1122
Db 3494 ATCGAGACACTCCCTGAGAAATACACACAGAGTAGGAGACAAAGGACCAAGCTCTCT 3553
QY GlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeu 1142
Db 3554 GGTGGCCAGAAACAGCGCATTTGCCATAGCTCGCGCTCTTGTGTAGACAGCTCATATTTG 3613
QY LeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAla 1162
Db 3614 CTTTGGATGAAGCTACATCAGCTCTGGATACAGAAAGTGAAGGTTGTCCAAAGAGCC 3673


```

Db 3314 GGCAAGAGATAAGCACCTGATGTCAGTGTCCGAGCACACCTGGCGCATCGTGTCT 3373
Qy 1063 GlnGluProValLeuPheAsnGlySerLeuAlaGluAsnLeuLeuValGlyAspAsnSer 1082
Db 3374 CAGGAGCCATCTCTGTTGACTGACGATTCGCGAAGCAATTCCTATGAGACAAACAGC 3433
Qy 1083 ArgValValProLeuAspGluLeuLeuGluAlaAlaAsnAlaAlaAsnLeuHisSerPhe 1102
Db 3434 CGGGTGTATCATGACATGAGAGATTATGAGGAGCCCAAGGAGGCCAATACACCACTTC 3493
Qy 1103 IleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSer 1122
Db 3494 ATCCAGACATCTCCCTGAGAAATACACACGAGTAGGAGACAAAGAGAAACCCAGCTCTCT 3553
Qy 1123 GlyGlyGlnLysGlnArgLeuAlaLeuAlaAlaGluLeuGlnLysProLysLeuLeu 1142
Db 3554 GGTGGCCAGAAACAGCGCATGTCATAGCTCGCGCTCTTGTATGACAGCTCTATATTTG 3613
Qy 1143 LeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAla 1162
Db 3614 CTTTGGATGAAGCTACATCAGCTCTGGATACAGAAAGTGAAAGGTTGTCCAAGAGCC 3673
Qy 1163 LeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaLeu 1182
Db 3674 CTGGACAAAGCCAGAGAGCCGACCTGATGTCATGATCGCCACCGCTTGTCCACCATC 3733
Qy 1183 GlnAsnAlaAspLeuLeuValLeuHisAsnGlyLysLeuGlnGlyThrHis 1202
Db 3734 CAGAATGCAGATTAAATAGTGGTGTTCAGAAATGGCAAGTCAAGGAGCATGCGACAT 3793
Qy 1203 GlnGluLeuLeuArgAsnArgAspLeuTyrPheLysLeuValAsnAlaGlnSer 1220
Db 3794 CAACAGCTGTGCCCAAGAGCACTATTTTCATGTCAGTGTCCAGGCT 3847

RESULT 14
AX391099
LOCUS AX391099 4646 bp DNA linear PAT 19-MAR-2002
DEFINITION Sequence 1 from Patent WO0210205.
ACCESSION AX391099
VERSION AX391099.1 GI:19584228
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Ruth,A. and Roninson,I.
Mutations of the mdri p-glycoprotein that improve its ability to
confer resistance to chemotherapeutic drugs
Patent: WO 0210205-A 1 07-FEB-2002;
THE BOARD OF TRUSTEES OF THE UNIVERSITY OF ILLINOIS (US)
FEATURES
source
    location/Qualifiers
    1..4646
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    425..4267
    /note="unnamed protein product"
    /codon_start=1
    /protein_id="CAD28599.1"
    /db_xref="GI:19584228"
    /translation="MDLEGDRNGAKKKFLLNKKSEKDKKKKPTVSVFMSFRYSN
WLDKLVVVGTLAAIITAGLPLMLVFGEMTDIIPANAGNLEDLMSNIINRSDINDTG
FFNNLEDMTRVAYYSYGAGLVAAIYQVSPWCLAAGRQIHKIKQFHFALMRQEI
GMPDHDVDELNTRLTDDVSKNEGIDGKIMFPQSMATFFTGIVGFRGKWLILVI
LAISPVLSAAVWAKILSFTDKELLAYAKAGAAVEALAIRTVIAGGQKKLEL
YNNLEAKRIGIKKAITANISIGAAFLIYASALAFWYGTTLVLSGYSIGQVLT
FRSLVIGATVSGASPISEAFANARGAAVEIFKIDNKPISIDSKSGHKPNKGNL
EFNVHFSYSPSRKEVKILGKLNLVQSGOTVALVGNSSGCKSTTVOLMORLYDPTEGM
VSDVDGRTINVRFLREIIGVVSQSPVLPATIAENIRYGRNVTMDSEIAKVEAN
AYDFIMKLPHKFTLVGERGAQUSGGQKQRIARALVRNPKILLDEADSEADTSE
AVQVQADKARKGRTITVIAHRLSTVRNADVAGPDGQVIVKEGNDELMLKKGIFFK
LVTMQTAGNEVELENAADESKSEIDALEMSSNDSRLIKRSTRSVRSGQAQDRKL

```

```

STKEALDESIPPVSWFRIMKLNLTWEPYFVVGVFCALINGLOPAPALIFSKILGVFT
RIDDPETKRONSNLFSLLFLALGIIISFITPFGQFTFGKAGEIILTKLRYVWFRSMLR
QDVSWDDPKNTGALTTRLANDAAQVKGATGSRVLAITONIANLGTGIIISFTYGWQ
LDTLLLAIVPIIAIAGVVMKMSQALKDKKELEGAGKATEAIENFRNTVVSQEQ
KFEHMYAQSLQVPYRNSLRKAHIFGITFSFTQAMMYFSYACGFRGAVLVAHKLMSFE
DVLVFSAVVFGAMAVGVQSSFADYAKAKISAHHILMIJEKTLIDSYSTEGMLPNT
LEGNVTGGEVFNYPTRPDIPVLOGLSLEVKGQTLALVSGSGCGKSTVQVLLERFYD
PLAGKLLLDGKEIKRIANVQMLRAHLGVISQEPILFDSCIAENIAGKDSNRVVSDEIV
RAAKEANTHAPIESLPNKYSTKGDKGTQSGQKQRIARALVRNPKILLDEADTSE
ALDTESEKVOEALDKAREGRTICIVIAHRLSTIQNADLIVVFQNGRVRKEHGTQOLLA
QKGIYFSMVSVQAGTKRQ"
BASE COUNT 1372 a 892 c 1130 g 1252 t
ORIGIN
Alignment Scores:
Pred. No.: 4,95e-227 Length: 4646
Score: 3530.50 Matches: 682
Percent Similarity: 74.00% Conservative: 246
Best Local Similarity: 54.39% Mismatches: 263
Query Match: 57.32% Indels: 63
DB: 6 Gaps: 10
US-09-873-409-5 (1-1222) x AX391099 (1-4646)
Qy 1 MetIleLeuGlyIleLeuLysSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 575 ATGTGGTGGGAACTTTGGCTGCCATCCATGGGCTGGAGCTTCTCTCATGATGCTG 634
Qy 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
Db 635 GTGTTGGAGAATGACAGATATCTTTCGCAATGCGAGGAATTTAGAGATCTGATGTC 694
Qy 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
Db 695 ACATCATCTAATAGAGTATATCATGATCAGGCTTCTTCAATCTGGAGGAGAC 754
Qy 44 -----ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuPheGlyTyr 61
Db 755 ATGACCAGGTATGCTTATTTATACAGTGGAAATTTGGTGTGGGTGCTGCTTAC 814
Qy 62 IleGlnIleSerLeuTyrIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db 815 ATTCAGGTTTCATTTTGGTGGCTGGAGCAATATACAAAATTTAGAAACAG 874
Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101
Db 875 TTTTTCATGCTATAATCGACAGGAGATAGCTGGTTGATGTCACCATGTTGGGAG 934
Qy 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db 935 CTTAACACCCGACTTACAGATGATGCTCTAAGATTAAATGAAGGAATTTGGTGACAAA 994
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db 995 GGAATGTTCTTTCAGTCAATGCGCAATTTTTCACCTGGGTTTATAGTAGATTACACGT 1054
Qy 141 GlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db 1055 GGTTGGAAGCTAACCTTGTGATTTGGCCATCAGTCTCTTCTTGGACTGTTCAGCTGCT 1114
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db 1115 GTCTGGGCAAGATACATATCTTCAATTTACTGATAAAGAACTCTTACGCTATGCAAA 1174
Qy 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db 1175 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACCTGTGATTCATTTGAGAGCAA 1234
Qy 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
Db 1235 AAGAAAGACTTTGAA-----PheValArgTyrThrGlnAsnLeuLysAsp 240
Qy 221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp

```



```
Qy 947 LysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysProAsnIleAspSerArg 966
|||
Db 3428 AAAATATCAGCAGCCCATCATCATCATGATCATGTAAGAAAACCCCTTGATTGACAGCTAC 3487
|||
Qy 967 SerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSer 986
|||
Db 3488 AGCAGCGAAGCCCTAATGCGCAACACATTCGGAAGAAATGTCACATTTGGTGAAGTTGTA 3547
|||
Qy 987 PhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGlu 1006
|||
Db 3548 TTCAACTATCCACCCAGCCGACATCCCATGCTTCAGGAGATGACCCCTGGAGGTAAG 3607
|||
Qy 1007 ArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGln 1026
|||
Db 3608 AAGGCCAGAGCTGGCTCTGGTGGGAGAGTGGCTGTGGGAAGACACAGTGGTCCAG 3667
|||
Qy 1027 LeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAla 1046
|||
Db 3668 CTCCTGAGCGGTTCTACGACCCCTTCGCGAGGAAAGTGTCTGTATGTCGCAAGAAATA 3727
|||
Qy 1047 LysGluLeuAsnValGlnTyrLeuArgSerGlnIleAlaIleValProGlnGluProVal 1066
|||
Db 3728 AAGCGACTGAATGTTCAGTGGCTCCGAGCACACCTGGGCGATCGTGTCCAGGAGCCCATC 3787
|||
Qy 1067 LeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValPro 1086
|||
Db 3788 CTGTTTACGTGACGATGCTGTGAGAACATTCCTATGAGACACACAGCCGGTGTGTCA 3847
|||
Qy 1087 LeuAspGluLeuLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeu 1106
|||
Db 3848 CAGGAGAGATCGTGAGGCGCAGCAAGAGGAGCCCAACATACATGCTTCATCGAGTCACTG 3907
|||
Qy 1107 ProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLys 1126
|||
Db 3908 CCTAATAAATATAGCATTAAGTAGGAGACAAAGAACTCAGCTCTCTGTGTGGCCAGAAA 3967
|||
Qy 1127 GlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGlu 1146
|||
Db 3968 CAAGCATTCGCATAGCTGTGCTGCTGTTAGACAGCTCATATTTTGTCTTTGGATGAA 4027
|||
Qy 1147 AlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAla 1166
|||
Db 4028 GCCACGTGAGTCTGGATACAGAAAGTGAAGAGTGTTCGAAGAGCCCTGGACAAAGCC 4087
|||
Qy 1167 ArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAsp 1186
|||
Db 4088 AGAAGAGCGCCGACCTGATTTGATTTGCTCACCCTGTCCACCATCCAGAAATGCAGAC 4147
|||
Qy 1187 LeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeu 1206
|||
Db 4148 TTATAGTGGTGTTCAGATGGCAGAGTCAAGGATCAAGGATGCGACGATCAGCAGCTGCTG 4207
|||
Qy 1207 ArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
|||
Db 4208 GCACAGAAAGGCATCTATTTTCAATGGTCACTGTCAGGCT 4249
|||

RESULT 15
AF269224 4045 bp DNA linear SYN 04-JUL-2000
LOCUS Synthetic construct Canis familiaris his-tagged-multidrug
DEFINITION resistance glycoprotein gene, complete cds.
ACCESSION AF269224
VERSION AF269224.1 GI:8926216
KEYWORDS synthetic construct.
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1 (bases 1 to 4045)
AUTHORS Roulet,A.C., Puel,O., Gesta,S., Alvinerie,M. and Pineau,T.J.
TITLE Analysis of dog MDR1 p-glycoprotein
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 4045)
```

AUTHORS
TITLE
JOURNAL

Roulet,A.C., Puel,O., Gesta,S., Alvinerie,M. and Pineau,T.J.
Direct Submission
Submitted (19-MAY-2000) Pharmacology, INRA, 180 Chamin de
Tournefeulle, BP3, Toulouse 31311, France

FEATURES
source

1..4045

/organism="synthetic construct"

/db_xref="taxon:32630"

/note="original mRNA isolated from Beagle dog colon"

101..3964

/codon_start=1

/transl_table=11

/product="his-tagged-multidrug resistance glycoprotein

MDR1"

/protein_id="AAF81747.1"

/db_xref="GI:8926217"

/translation="MHHHHDDPEGGRKGSAAEKNFWMKMGKSKKNEKKEKKTPTVSTFA
MFRYSNWLDRLYMLVGTWAAIIHGALPLMLVFCNMTDSPANAGISRNKTPPVLINE
SITNTQHFHLEEMTYAYYSIGGAGVLVAAYIQVSPHCLAGRQILKIRQFF
HAIHQEIGFWDVHDVDELNRLTDDSKINEGIDGKIMFPQSIATFTGIVGFTF
GOKLTLVLAISPVLGSAAIWAKLSFTDKELLYAKAGAVAEVLAARTVAFG
GOKLELYNKLEAKRIGIKKAITANISIGAAFLIYASVALAFWGTSTLVLSSEY
STGOVLTVFFSVLIGAFSPSIEAFANARGAAEYFKIIDNKPISDSKSGHK
PDNIKGNLEFNHVESYSPRKVKILKGLNKVQSGQTVVALVNSGCGKSTVOLMOR
LIDPDGMWIDGQDIRIINVRHLREITGVVSQEPVLEATTIAENIRYRENTWDEI
EKAVKEANAYDFIMKLPNKFPDPLVGERGAQSGGQKQRIATIALVRNPKILLDEAT
SALDTSEAVVQVALDKARGTITVIAHRLSTVRNADVIAGFDGVIKEGNHDELM
KEGIVFKLVTMOTRNEIELENATGESKSDALEMSPKDSGLSIRRRSIIHA
POGQDRKLTGEDLNENVPVSWRILKLNSTEDPVPVYFVIFCAIINGLQPAFSIIF
SRIGIGTRDEDPETKRONSNMFSVLVLVGLIISFITPLOGFTPKAGEIILKRLY
MYFRMLKQDVSWDFDPPKNTGALTTRLANDAAQVKAIGSLAVITQNIANLGTGII
ISLIYQWQLTLLLAIVPIIAIAGVEMKUSQALKKKKELEGAGKIAELENPRT
VVSLEQREKFEYMAQSLQVPMVNSRLKAHIFGVFSITQMMYFIACGFRFGLYL
ANEFNFDVLLVFSAIIVFGMAVQSVSPADYAKAKVSAAHVIMIIEKSPGLDYS
PHLKPNTLEGVNTNEVFNPTPDIPVLQGLSLEVKKGOTLALVSGSGCGKSTV
QLLERPYDPLAGSVLIDGKELKHLNOMLRAHLGIVSOEPIIFDCSIAENIAYGDNRS
VVSHEIWAKEANHHIFETLPEKYNTRVGDKGTQSGGQKQRIATIALVRQPHI
LLSDTSALDTESEKVOEALDKAREGRTCVIAHRLSTIQNADLIIVFQNGKVAEH
GTHQQLLAQKGIYFISMISVQAKR"

BASE COUNT 1171 a 850 c 986 g 1038 t
ORIGIN

Alignment Scores:

Pred. No.: 5,64e-227 Length: 4045
Score: 3528.50 Matches: 681
Percent Similarity: 73.13% Conservative: 239
Best Local Similarity: 54.13% Mismatches: 269
Query Match: 57.29% Indels: 69
DB: 12 Gaps: 9

US-09-873-409-5 (1-1222) x AF269224 (1-4045)

Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20

Db 269 ATGTGTGGTGGGACAATGGCTGCCATCATCCATGGAGCTGCATCCCTCTCATGATGCTG 328

Qy 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu----- 34

Db 329 GTTTTGGAAACATGACAGATAGCTTTGCAATGTCAGGAATTTCAAGAAACAAAACTTTT 388

Qy 35 -----ValGlnThrAsnThrTyrSerPhePheArg----- 44

Db 389 CCAGTTATAATTAATGAAGTATTACCAACAATACACACATTTTCATCAACCATCTCGAG 448

Qy 45 -----LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe 59

Db 449 GAGGAATGACACGATGCTTATTATACGTGGGATCGGTGCTGGCTGCTGGTCTGCTGCTG 508

Qy 60 GlyTyrIleGlnIleSerLeuTyrPheIleThrAlaAlaArgGlnThrLysArgIleArg 79

Db 509 GCTTACATCCAGGTTTCATTCTGTGCTGCTGCGAGCAGGAAGACAGATACATCAAAATAGA 568

Qy 80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIle 99

Db 569 AAACAATTTTTTCATCTATCATCGACAGGAGATTGGCTGTGTTACGTGCATGACGTT 628
Qy GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
Db 629 GGGGAGCTTAAACACCGGCTCACAGACGATGCTCAAAATCAATGAAGAAATGGCGAC 688
Qy LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
Db 689 AAAATTGGAATGTTCTTTCAATCAATAGCAACATTTTCCACCGTTTATATAGTGGGTTT 748
Qy ValLysGlyTTPlysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
Db 749 ACACGTGGTGAAGCTAACCCCTGTGATTTTGGCCATCAGCCCTGTTTGGCACTTCA 808
Qy 159 AlaAlaAlaCysSerArgMetValIleSerLeuSerLysGluLeuSerAlaTyrSer 178
Db 809 GCGCCCATCTGGCAAGATACCTATCTTCAATTACTGATAAAGAACTCTTGGCTATGCA 868
Qy LysAlaGlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArg 198
Db 869 AAAGCTGGAGCAGTACGTGAAGAAGTCTTAGCAGCAATCAGAACTGTGATTTGCTTGG 928
Qy 199 AlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPhe 218
Db 929 GGCACAAAGAAAGAACTTGAA----- 949
Qy 219 TyrPheProGlnTrpLeuLeuSerCysValLeu**PheValArgTyrThrGlnAsnLeu 238
Db 950 -----AGGTACAAACAAAATTTA 967
Qy 239 LysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAla 258
Db 968 GAAGAAGCTAAAGAATTTGGATTAAGAAAGCTATCACGSCCAACATTTCTATTGGTGC 1027
Qy 259 ValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIle 278
Db 1028 GCCTTCTTATGATCATGTCATCATGCTCTGGCTTCTGGTATGGGACCTCTTGGTC 1087
Qy 279 LeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIle 298
Db 1088 CTCTCCAGTGAA-----TATTCTATGGCAAGTACTCACTGCTCTTCTTCTGATTATA 1141
Qy 299 HisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArg 318
Db 1142 ARTGGGCTTTAGTATTGGACAGGATCCCAAGCATTTGAACCATTTGCAACGCAAGA 1201
Qy 319 GlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSer 338
Db 1202 GGAGCAGCTATTGAATCTTCAAGATAATTGCAATAAACAAGCATTTGACAGCTATTGC 1261
Qy 339 ThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPhe 358
Db 1262 AAGAGTGGACATAAACAGATAATATAAGGGAAATTTGGAAATTCAAAATGTTCACTTC 1321
Qy 359 AsnTyrProSerArgProSerIleLysIleLysGlyLeuAsnLeuArgIleLysSer 378
Db 1322 AGTTACCTTCTCGAAAAGAGTTAAGATCTTAAAGGTTCTCAACCTGAAGTTTCAGAGT 1381
Qy 379 GlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeu 398
Db 1382 GGGCAGACAGTGGCGTGTGGGAACAGTGGCTCGGGAAGACGACGACCGTGCAGCTG 1441
Qy 399 LeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArg 418
Db 1442 ATGCAGAGGCTCTATGACCCACAGATGGCATGGTCTGTATTGGACAGGACATTAGG 1501
Qy 419 AlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeu 438
Db 1502 ACCATAAATGTAAGGCATCTTCGGGAAATTAATTCTGGTGTGTGAGTCAGGAGCCTGTGTTG 1561
Qy 439 PheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspValThrAspGluGlu 458
Db 1561 ----- 469

Db 1562 TTTGCCACCAACATAGCTGAAACCAATTGCTATGCGCGCGAAAAATGTCACCATGGATGAG 1621
Qy 459 MetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProLeuLys 478
Db 1622 ATTGAGAAAGCTGTTAAGGAAGCCCAATGCTATGATTTTATCATGAACACTACCTAATAAA 1681
Qy 479 PheAsnThrLeuValGlyGlyLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIle 498
Db 1682 TTTGACACTCTGTTGGAGAGAGAGGGGCCAGCTGAGTGGTGGACAGAAACAGAAATC 1741
Qy 499 AlaIleAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSer 518
Db 1742 GCATTTGCTGGGCCCTGTTCCGAACCCCAAGATTTCTTCTGCTGGATGGCAACGTC 1801
Qy 519 AlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGly 538
Db 1802 GCTCTGGACACTGAAAGTGAAGCAGTGGTTTCAAGTGGCCCTGGATAAAGCCAGAAAGGC 1861
Qy 539 ArgThrThrIleValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleVal 558
Db 1862 CGGACTACCATTTGTGATAGCTCATGCTTTGTCTACAGTTTCTGAATGCCATGTCTATGCT 1921
Qy 559 ThrLeuLysAspGlyMetLeuAlaGlyLysGlyAlaHisAlaGluLeuMetAlaLysArg 578
Db 1922 GGTTTTGATGAGTGGATCATTTGTGGAGAAAGGAATCATGATGAACCTCATGAAGAAG 1981
Qy 579 GlyLeuTyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGlu 598
Db 1982 GGCATTACTTCAAACTTGTCAATGTCAG---ACAAGAGGAAATGAAATTTAGTTAGAA 2038
Qy 599 SerMetTyrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSer 618
Db 2039 AATGCCACTCGTGAATCCAAAAGTGAAGTGAAGTATGCTTGGAAATG---TCTCCAAAGAT 2095
Qy 619 IleLysSerAspPheIle----- 624
Db 2096 TCAGGCTCCAGTTTATAAAAAAGAGATCACTCGCAGGAGTATACATGCACCACCAAGGC 2155
Qy 625 ---AspLysAlaGluLeuSerThrGlnSerLysGluIleSerLeuProGluValSerLeu 643
Db 2156 CAAGACAGAAAGCTTGGTACAAAAGAGGACTTGAATGAGAATGCTACCTCCAGTTTCTTTC 2215
Qy 644 LeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAla 663
Db 2216 TGAGGATTTCTGAAGCTGAACCTCAACTGAATGCGCTTATTTTGTGGTGTATATTTGT 2275
Qy 664 SerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThr 683
Db 2276 GCTATTATAACCGGAGCTCGCAACCAAGCATTTTCAATAATATTTTCAAGGATTTATAGG 2335
Qy 684 MetPhe---GlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMet 702
Db 2336 ATCTTTACCGAGATGAGGATCTTGAAACAAAACGACAGATAGTAAACATGTTTCTGTA 2395
Qy 703 IlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyr 722
Db 2396 TTGTTTCTAGTCTTGAATTTATTTTATTTATACATTTTCTCCAGGCTTTCACATTT 2455
Qy 723 GlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeu 742
Db 2456 GGCAAAGCTGGGAGATCTCACTAAGCGGCTTCGATATGATGGTTTTCAGATCCATGCTG 2515
Qy 743 TyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIle 762
Db 2516 AGCAGATGTCAGCTGGTTTGTATGATCCCTTAAACACCACTGGAGCATTGACACAGG 2575
Qy 763 LeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThr 782
Db 2576 CTTCGCAATGATCGGCTCAAGTTAAAGGGCTATAGTTTCCAGGCTTGTGCTCATTACC 2635
Qy 783 GlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMet 802
Db 2636 CAGATATAGCAAACTCTGGGACAGGCACTTATATATCTTATCTTATCTTGTGTCGAATTA 2695

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 31, 2003, 13:52:37 ; Search time 434.971 Seconds
(without alignments)
2337.888 Million cell updates/sec

Title: US-09-873-409-6
Perfect score: 6012
Sequence: 1 MILGILASLVNACGLPLMPL.....QELLNRDIYKLVNAOSVQ 1195

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 574371 seqs, 425486471 residues

Total number of hits satisfying chosen parameters: 1148742

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n_model -DEV=xlh
-Q=/cgn2_1/USPTO_spo3/US09873409/runat_27032003_115424_19290/app_query.fasta_1.7544
-DB=Published Applications NA -Qfmt=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -SPART=1 -END=-1 -MATRIX=blom62
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09873409 @CCN 1.1 960 @runat_27032003_115424_19290
-NCPU=6 -ICPU=3 -NO_XLPXY -NO_MAP -LARGEQUERY -NEG_SCORES=0 WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA.*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
- 14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	6012	100.0	3621	US-09-873-409-14
2	5988.5	99.6	3702	US-09-873-409-13
3	5159.5	85.8	3177	US-09-873-409-12
4	4561.5	75.9	2856	US-09-873-409-10

5	3554	59.1	3860	10	US-09-866-866A-1	Sequence 1, Appli
6	3549	59.0	4317	9	US-10-044-671-1	Sequence 1, Appli
7	3545	59.0	3860	10	US-09-866-866A-3	Sequence 5, Appli
8	3545	59.0	8630	10	US-09-306-417-1	Sequence 1, Appli
9	3545	59.0	8630	10	US-09-306-417-2	Sequence 2, Appli
10	3540.5	58.9	4425	10	US-09-769-097-3	Sequence 3, Appli
11	3538.5	58.9	4369	10	US-09-769-097-1	Sequence 1, Appli
12	3537.5	58.8	4643	9	US-10-072-621-2	Sequence 2, Appli
13	3526.5	58.7	4189	10	US-09-866-866A-5	Sequence 5, Appli
14	3522.5	58.6	4788	10	US-09-866-866A-7	Sequence 7, Appli
15	3504.5	58.3	3924	10	US-09-880-107-2299	Sequence 2299, Ap
16	3476.5	57.8	4254	10	US-09-917-800A-1424	Sequence 1424, Ap
17	3456	57.5	2066	10	US-09-873-409-9	Sequence 9, Appli
18	3428.5	57.0	3912	10	US-09-917-800A-1560	Sequence 1560, Ap
19	2835	47.2	5010	10	US-09-917-800A-483	Sequence 483, App
20	2552	42.4	1940	10	US-09-873-409-16	Sequence 16, Appl
21	2528.5	42.1	2021	10	US-09-873-409-15	Sequence 15, Appl
22	2380.5	39.6	4175	10	US-09-749-340-3	Sequence 3, Appli
23	2351.5	39.1	3861	9	US-09-938-842A-2237	Sequence 2237, Ap
24	2282.5	38.0	4002	9	US-09-758-828-1	Sequence 1, Appli
25	2282.5	38.0	4002	9	US-09-758-828-3	Sequence 3, Appli
26	2153	35.8	3861	9	US-09-938-842A-263	Sequence 263, App
27	2046	34.0	3512	10	US-09-749-340-7	Sequence 7, Appli
28	2010.5	33.4	4653	9	US-10-101-388-2	Sequence 2, Appli
29	1659.5	27.6	3792	9	US-09-882-694-10	Sequence 10, Appl
30	1617	26.9	2698	10	US-09-749-340-5	Sequence 5, Appli
31	1550	25.8	8036	9	US-10-101-388-1	Sequence 1, Appli
32	1538.5	25.6	2681	10	US-09-749-340-8	Sequence 8, Appli
33	1525	25.4	3999	9	US-09-882-694-9	Sequence 9, Appli
34	1102.5	18.3	1810	10	US-09-749-340-4	Sequence 4, Appli
35	1101.5	18.3	1175	10	US-09-873-409-11	Sequence 11, Appl
36	1084	18.0	7296	10	US-09-070-927A-59	Sequence 59, Appl
37	1031.5	17.2	3047	9	US-09-738-626-1251	Sequence 1251, Ap
38	1021	17.0	9747	10	US-09-070-927A-12	Sequence 12, Appl
39	1005	16.7	12438	10	US-09-070-927A-173	Sequence 173, App
40	941	15.7	64681	10	US-09-750-988-1	Sequence 1, Appli
41	902.5	15.0	3601	10	US-09-070-927A-255	Sequence 255, App
42	901.5	15.0	6415	10	US-09-070-927A-275	Sequence 275, App
43	781	13.0	2298	9	US-10-156-239-18	Sequence 18, Appl
44	781	13.0	2298	10	US-09-795-693-18	Sequence 18, Appl
45	781	13.0	3408	9	US-10-156-239-16	Sequence 16, Appl

ALIGNMENTS

RESULT 1
US-09-873-409-14
; Sequence 14, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
; APPLICANT: Frank, Markus
; APPLICANT: Sayegh, Mohamed
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873,409
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
; LENGTH: 3621
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-873-409-14

Alignment Scores:
Pred. No.: 0
Score: 6012.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 10
Length: 3621
Matches: 1195
Conservative: 0
Mismatches: 0
Indels: 0
Gaps: 0

US-09-873-409-6 (1-1195) x US-09-873-409-14 (1-3621)

QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
DB 34 ATGATCTGGGTATACCTGGCATCACTGGTCAATGGAGCCTGCCTTCCTTTAAATGCCACTG 93
QY 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
DB 94 GTTTTAGAGAAATAGTGATTAACCTTATTAGTGGATGTCTAGTCCAAACTAAACATAC 153
QY 41 SerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaIleLeuPheGly 60
DB 154 TCCTTCCTCAGGTGACCCCTGATATTATGTGGATAGGTGTGCTGCCTTGATTTTGGT 213
QY 61 TyIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrIleAspIleArgLys 80
DB 214 TACATACAGATTTCCTTGTGGATTATTAACCTGCAGCAGACAGACCAGAGGATTCGAAAA 273
QY 81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGly 100
DB 274 CAGTTTTTTTCATTCACTGGTGGCACAGGACATCGGCTGGTTTGATAGCTGTGCATCGGT 333
QY 101 GluLeuAsnThrArgMetThrAspIleAspIleSerAspGlyIleGlyAspLysIle 120
DB 334 GAACCTTAACACTCGCATGCAGACATTCACAAAATCAAGTATGAGATTAAGATT 393
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
DB 394 GCTCTGTGTGTTTCAAAACATGTCTACTTTTTCGATTGGCTGGCAGTTGGTTGGTGAAG 453
QY 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
DB 454 GGCTGGAAACTCACCTTAGTGACTCTATCACAGTCTCTCTTATATATGGCTTCAGGGCA 513
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
DB 514 GCATGTCTTAGGATGGTCATCTCATTCACAGTAAGGAATTAAGTGCCTATTCCAAAGCT 573
QY 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
DB 574 GGGCTGTGCAGNAGAGTCTTGTCATCAATCCGAACAGTCATAGCTTTTAGGGCCAG 633
QY 201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
DB 634 GAGAAAGAACTTCAAGGTATACACAGATCTCAAGATGCAAAAGGATTTTGGCATAAAA 693
QY 221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
DB 694 AGGACTATAGCTTCAAAAGTGTCTCTGGTGTGTGTACTTCTTTATGAATGGAACCTAT 753
QY 241 GlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
DB 754 GGACTTGCTTTTGGTATGNAACCTCTGATCTTAAATGGAGNACCTGGATATACCATC 813
QY 261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
DB 814 GGGACTGTCTCTGCTGTTCTTTAGTGTATCCATAGCAGTTATTGCAATGGAGCAGCA 873
QY 281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
DB 874 GTCCCTCACCTTTGAAACCTTCGCAATAGCCGAGAGCTGCCTTTCATATATTTCACGGTT 933
QY 301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
DB 934 ATTGATGAAGAAACCCAGTATAGATAACTTTTCCACAGCTGGATATAAACCTGGAATCCATA 993
QY 321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
DB 994 GAAGGAACCTGTGGAATTTAAAAATGTTCTTCAATTAATCCATCAAGACCATCTATCAAG 1053
QY 341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
DB 1054 ATTTCTGAAGGCTCTGAATCTCAGAAATTAAGTCTGGAGAGACAGTCGCTTGGTGGTCTC 1113

QY 361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
DB 1114 AATGGCAGTGGGAAGAGTAGTACGGTAGTCCAGCTCTTCAGAGGTTATATGATCCGGATGAT 1173
QY 381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
DB 1174 GGCTTTTATCATGTGTGATGAGAATGACATCAGAGCTTTTAAATGTGCGCATTTATCGAGAC 1233
QY 401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
DB 1234 CATATTGGAGTGTGTAGTCAAGAGCCTGTTTGTTCGGACCCACCATCAGTAAACATATC 1293
QY 421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
DB 1294 AACTATGGACGAGATGATGTGACTGATGAAGAGATGGAGAGACAGCAAGGGAAGCAAT 1353
QY 441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
DB 1354 GCGTATGATTTTATCATGTGAGTTTCCATAAATTTAATACATTTGGTAGGGGAAAAAGGA 1413
QY 461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
DB 1414 GCTCAATGAGTGGAGGGCAGAAACAGAGATCGCAATTTGCTCGTGCCTTAGTTCGAAC 1473
QY 481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
DB 1474 CCCAAGATTCTGATTTTAGATGAGGCTACGTCTGCCCTGGATTTCAGAAAGCAAGTCAGCT 1533
QY 501 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 520
DB 1534 GTTCAAGCTGCATCGGAAGGCGAGCAAAAGTTCGAGCTACAAATCGTGTGTAGCACACCGA 1593
QY 521 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 540
DB 1594 CTTTCTACTATTTCGAAGTGCAGATTTGATTGTGACCCTTAAAGGATGGAATGCTGGCGGAG 1653
QY 541 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 560
DB 1654 AAAGGACACATGCTCAACTAATGCAAAACAGAGGTCTATATTATTACCTTGTGATGTCA 1713
QY 561 GlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLys 580
DB 1714 CAGGATATTAAAAAAGCTGATGAACAGATGGAGTCAATGACATATTTCTACTGAAGAAAG 1773
QY 581 ThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAla 600
DB 1774 ACCAATCACTTCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTCATTGACAGGCT 1833
QY 601 GluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeu 620
DB 1834 GAGGAATCCACCCCAATCAAGAGATAAGTCTTCTCTGAAGTCTCTCTATAAAAAATTTTA 1893
QY 621 LysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsn 640
DB 1894 AAGTTAAACAAGCCTCAATGGCCTTTTGTGGTTCTGGGACATTTGGCTTCTGTTCTAAAT 1953
QY 641 GlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn 660
DB 1954 GGAACGTGTTTCATCCAGTATTTCATCATCTTTTGCAGAAATTTATAACCATGTTTGGAAAT 2013
QY 661 AsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeu 680
DB 2014 AATGATATAAACCATTAAGCATGATGCAGAAATTTATTCATGATATTTCGTCATTTCG 2073
QY 681 GlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGlu 700
DB 2074 GGTGTTATTGCTTCTCAGTTATTTCATGCAGGGATTTATTTACGCGCAGACAGCGGAA 2133
QY 701 IleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAla 720
DB 2134 ATTTTAACGATGAGATTAAGACACTTGGCCCTTCAAAGCCCAATGTTATATCAGGATATTGCC 2193

QY 721 TrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIle 740
DB 2194 TGGTTTGGATGAAAGGAAACACGACAGGAGGCTTGACAAACATATTAGCCATAGATATA 2253
QY 741 AlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsn 760
DB 2254 GCACAAATTCAGAGGACACAGGTTCCAGATTGGCGTCTTAACACAAATGCAACTAAC 2313
QY 761 MetGlyLeuSerValIleIleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeu 780
DB 2314 ATGGGACTTTCAGTTATCATTTCTTTATATATGGATGGGAGATGACATTCCTGATTCTG 2373
QY 781 SerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPhe 800
DB 2374 AGTATTGCTCCAGTACTTGCCTGCAGAGGAATGATTGAAACCCGACGAATGACTGGATT 2433
QY 801 AlaAsnLysAspLysGlnGluLeuLysHisAlaGlyIleAlaThrGluAlaLeuGlu 820
DB 2434 GCCAACAAAGATAAGCAAGAACTTAGACATGCTGGAAAGATAGCAACTGAAGCTTTGGAG 2493
QY 821 AsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGlu 840
DB 2494 AATATACGTACTATATGTCATTTACACAGGAAAGCCCTTCAGGCAATGTATGAAGAG 2553
QY 841 MetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyr 860
DB 2554 ATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCACAGATATTGGAAGCTGTAT 2613
QY 861 AlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyr 880
DB 2614 GCATTGAGCCATGCCCTTTATATATTTTGGCTATGACAGAGGGTTTCGATTGGAGGCTAT 2673
QY 881 LeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAla 900
DB 2674 TTAATTCAGCTGGACGATGACCCAGAGGGCATGTTTCATAGTTTACTGCAATGCA 2733
QY 901 TyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLys 920
DB 2734 TATGGAGCTATGGCCATCGAAACCGCTCGTTTGGCTCCTGAATATTCCAAAGCCAAA 2793
QY 921 SerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSer 940
DB 2794 TCGGGGGCTGCGCATCTGTTTGGCTTGTGGAAAGAAACCAATATAGACAGCCGACGT 2853
QY 941 GlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhe 960
DB 2854 CAAGAAGGAAAGCCAGACACATGTGAAGGAAATTTAGAGTTTCAGAAAGTCTCTTTC 2913
QY 961 PheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArg 980
DB 2914 TTCTATCCATGTCGCCAGATGTTTTCATCTCCGTGGCTTATCCCTCAGTATTGAGCGA 2973
QY 981 GlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeu 1000
DB 2974 GGAAAGACAGTAGCATTTTGTGGGAGCAGCGGCTGTGGGAAAGACATCTCTGTCAACTT 3033
QY 1001 LeuGlnArgLeuTyrAspProValGlnGlnValLeuPheAspGlyValAspAlaLys 1020
DB 3034 CTCAGAGACTTTATGACCCCGTGACAGGACAGTGTCTGTTGATGGTGGAGTGCAGAAA 3093
QY 1021 GluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeu 1040
DB 3094 GAATTGAATGTACAGTGGCTCCGTTCCCAATAGCAATCGTTCTCAAGAGCCTGTGCTC 3153
QY 1041 PheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeu 1060
DB 3154 TTCAACTGTCAGCATTTGTCAGAACATCGCTATGTTGACCAACAGCCGCTGTGGTGCCATTA 3213
QY 1061 AspGluIleLysGluAlaAlaAsnAlaAsnIleHisSerPheIleGluGlyLeuPro 1080
DB 3214 GATGAGATCAAGAAGCCGCAATGACGCAAAATATCCATTTCTTTATTGAAGGTCTCCCT 3273
QY 1081 GluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGln 1100

DB 3274 GAGAATAACAACACAGTTGGACTGAAAGGAGGACACAGCTTTCTGGCGCCAGAAACAA 3333
QY 1101 ArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAla 1120
DB 3334 AGACTAGCTATTGCAAGGCTCTTCTCCAAAACCCAAAATTTTATGTTGGATGAGGCC 3393
QY 1121 ThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArg 1140
DB 3394 ACTTCAGCCCTCATATAATGACAGTGAGAAAGTGGTTTCAGCATGCCCTTGATAAGCCAGG 3453
QY 1141 ThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeu 1160
DB 3454 ACGGGAGAGACATGCTAGTGTCTCACTCAGGCTCTCTGCAATTTCAGAACCGCAGATTG 3513
QY 1161 IleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArg 1180
DB 3514 ATATGTTCTCGCANATGGAAGATTAAGGAACAAGGAACATCATCAAGAGCTCTTGAGA 3573
QY 1181 AsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 1195
DB 3574 AATCGAGACATATATTTAAGTTAGTGAATGCACAGTCAGTGCAG 3618

RESULT 2

US-09-873-409-13
; Sequence 13, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
; APPLICANT: Frank, Markus
; APPLICANT: Sayegh, Mohamed
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873,409
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 13
; LENGTH: 3702
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: NO. US20020037522A1e
; LOCATION: (723)..(723)
; OTHER INFORMATION: n at position 723 represents any nucleotide (A, T, C or G)
US-09-873-409-13

Alignment Scores:
Pred. No.: 0 Length: 3702
Score: 5988.50 Matches: 1195
Percent Similarity: 97.79% Conservative: 0
Best Local Similarity: 97.79% Mismatches: 0
Query Match: 99.61% Indels: 27
DB: 10 Gaps: 1

US-09-873-409-6 (1-1195) x US-09-873-409-13 (1-3702)

QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
DB 34 ATGATCCTGGTATATCTGGCATCCTGGTCAATGGAGCCCTTCTTTAATGCCACTG 93
QY 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
DB 94 GTTTTAGGAGAAATGATGATAACCTTATTAGTGGATGCTAGTCCAAACATAACATAC 153
QY 41 SerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGly 60
DB 154 TCTTTCTTCAGGTGACCTCTATTATTATGTAATAGGTGTTGCTGCTGCTTTTGGT 213
QY 61 TyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLys 80
DB 214 TACATACAGATTTCTTGTGGATTATTAACGTGACGACGACAGAGGATTCGAAAA 273

QY 81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGly 100
DB 274 CAGTTTTTTCATTTCAGTTTTGGCACAGGACATCGGCTGGTTTGATAGCTGTGACATCGGT 333
QY 101 GluLeuAsnThrArgMetThrAspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
DB 334 GAATTAACCACTCGCATGACAGACATTTGACAAATCAGTGATGGTATGAGATAAGATT 393
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
DB 394 GCTCTGTGTGTTTCAAAACATGTCTACTTTTTCGATTTGGCTGGCAGTTGGTGGTGAAG 453
QY 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
DB 454 GGCTGGAACTCACCTAGTACTCTATCCAGCTCTCTCTTATATGGCTTCAGCGGCA 513
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTrpSerLysAla 180
DB 514 GCATGTGTAGGATGGTCACTCTATTGACCAGTAAGGAATTAAGTGCTTATTCCTCAAGCT 573
QY 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
DB 574 GGGCTGTGGCAAGAAAGTCTTGTCTATCAATCCGAACAGTCATAGCCITTTAGGGCCAG 633
QY 201 GluLysGluLeuGln----- 205
DB 634 GAGAAAGAACTTCAAGGCTCTTCCITTTAAATATAACAAGATATGCTGGTTTATTTT 693
QY 206 -----ArgTyrThrGlnAsnLeuLysAsp 213
DB 694 CCCAGTGGCTACTAAGTTGTTCTCTGTTTGTAAAGGTATACACAGAATCTCAAGAT 753
QY 214 AlaLysPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 233
DB 754 GCAAGGATTTGGCATAAAGAGACTATAGCTTCAAAAGTGTCTCTGTGTGTGTGTAC 813
QY 234 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 253
DB 814 TTCTTTATGATGGAACCTATGGACTTGTCTTTTGGTATGGAACCTCTTGATCTTAAT 873
QY 254 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 273
DB 874 GGAAACCTCGATATACATCGGAGCTGTCTTCTGCTGTTTCTTTAGTGTAAATCCATAGC 933
QY 274 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 293
DB 934 AGTTATTCATTTGGAGCAGCAGTCCCTCACTTTGAAACCTTCGCAATAGCCCGAGGAGCT 993
QY 294 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 313
DB 994 GCCTTTTCATATTTCCAGGTATTTGATAGAAACCCAGTATAGATAACTTTTCCACAGCT 1053
QY 314 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 333
DB 1054 GGATATAAACCTGAATCCATAGAGGAACCTGTGGAATTTAAATAATGTTCTTCAATAT 1113
QY 334 ProSerArgProSerIleLysIleLysGlyLeuAsnLeuArgIleLysSerGlyGlu 353
DB 1114 CCATCAAGACCATCTATCAAGATTCTCAAGAGTCTGGAATTTCAAGATTAAGTCTGGAGAG 1173
QY 354 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 373
DB 1174 ACATCGCCTTGGTCTCTCAATGGCAGTGGGAGAGTACGGTAGTCCAGCTTCTGCAG 1233
QY 374 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 393
DB 1234 AGGTTATATGATCCGGATGATGGCTTTTATCATGGTGGATGAGAATGACATCAGAGCTTTA 1293
QY 394 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 413
DB 1294 AATGTGGGCAATATTCAGACCATATTTGGAGTGGTGTAGTCAAGAGCCTGTTTGTTCGGG 1353
QY 414 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspValThrAspGluGluMetGlu 433

DB 1354 ACCACCATCAGTAACAATATCAAGTATGACGAGATGATGTGCTACTGATGAAGAGATGGAG 1413
QY 434 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 453
DB 1414 AGAGCAGCAAGGGAAGCAAAATCGGTATGATTTTATCATGGAGTTTCTTAATAAATTAAT 1473
QY 454 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 473
DB 1474 ACATTTGGTAGGGGAAAAAGGAGCTCAATAGTGGAGGCGACAAACAGAGATTCGCAAT 1533
QY 474 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 493
DB 1534 GCTCGTGCCTTAGTTGCAAAACCCCAAGATCTTGAATTTAGATGAGGTACGTCTGCCCTG 1593
QY 494 AspSerGluSerLysSerAlaValGlnAlaLeuGluLysAlaSerLysGlyArgThr 513
DB 1594 GATTCAGAAAGCAAGTCAGCTGTTCAAGCTGCAGTGGAGAGCGCAAGGTCGGACT 1653
QY 514 ThrIleValAlaAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 533
DB 1654 ACATCTGTGTAGCACACCGACTTCTACTATTCGAAAGTGCAGATTTGATTGTGACCTTA 1713
QY 534 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 553
DB 1714 AAGGATGGAATGCTGGCGGAGAAAGGAGCACATGCTGAATTAATGGCAAAACGAGGTCTA 1773
QY 554 TyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMet 573
DB 1774 TATATTTCATCTGTGATGTGCAGAGGATATTTAAAGCTGTGATGAACAGATGGAGTCAATG 1833
QY 574 ThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLys 593
DB 1834 ACATATTCTACTGAAAGAAAGACCAACTCACTCTCTGCACCTCTGTGAAGACATCAAG 1893
QY 594 SerAspPheIleAspLysAlaGluSerThrGlnSerLysGluLysSerLeuProGlu 613
DB 1894 TCAGACTTCATTGACAAGCTCGAGGAATCCACCAATCTAAAGAGATAAGTCTTCCTGAA 1953
QY 614 ValSerLeuLeuLysIleLysLeuAsnLysProGluTrpProPheValValLeuGly 633
DB 1954 GTCTCTCTATTAATAAATTTTAAAGTTAAACCAAGCTGGAATGGCCTTTTGTGGTCTGGGG 2013
QY 634 ThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLys 653
DB 2014 ACATTCGCTTCTGTTCTTAATGGAACCTGTTTCATCCAGTATTTTCCATCATCTTTGCAAAA 2073
QY 654 IleIleThrMetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyr 673
DB 2074 ATTATAACCATGTTTGGAAAAATATGATAAAACCAATTAAGCATGATGCAGAAATTTAT 2133
QY 674 SerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeu 693
DB 2134 TCCATGATATTCGTCAATTTGGGTGTATTTGCTTTGTTCAGTTATTTTCATGCGAGGATTA 2193
QY 694 PheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAla 713
DB 2194 TTTTACGGCAGACGAGGGGAAATTTTAAAGTAGATTAAGACACTTGGCCTTCAAGGCC 2253
QY 714 MetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThr 733
DB 2254 ATGTTATATCAGGATATTCCTGCTTGTGTTTGAAGGAAAAACAGCAGAGGCTTGACA 2313
QY 734 ThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyVal 753
DB 2314 ACAATATTAGCCATAGATATAGCACAAATTCAGAGGAGCAACAGGTTCAGGATTTGGCGTC 2373
QY 754 LeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyr 773
DB 2374 TTAACACAAATTCACACTAACATGGGACTTTTCAGTTATCATTTCTTTATATATGATGG 2433
QY 774 GluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGlu 793

Db 2434 GAGATGACATTCCTCGATTCTGAGTATTGCTCCGATGACTTCCCGTGACAGCAATGATTGAA 2493
Qy 794 ThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLys 813
Db 2494 ACCGACGCAATGACTGGATTTCGCAACAAGATAAGCAAGAACTTAAGCATGCTGGAAG 2553
Qy 814 IleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAla 833
Db 2554 ATAGCAACTGGAAGCTTTGGAGAAATATACGTACTATATAGTGTCTATTAAACAAGGGAAAAAGCC 2613
Qy 834 PheGluGlnMetYrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAla 853
Db 2614 TTGAGCAAAATGATGAGAGATGCTTTCAGACTCAACACAGAAATACCTCGAAGAAGCA 2673
Qy 854 GlnIleGlySerCysYrAlaPheSerHisAlaPheIleYrPheAlaYrAlaAla 873
Db 2674 CAGATTATTGGAAGCTGTTATGCAATTCAGCCATGCTTTATATATATTGCTATGACGA 2733
Qy 874 GlyPheArgPheGlyAlaYrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPhe 893
Db 2734 GGGTTTCGATTTCGAGCCTATTTAATTCAGCTGGACGAATGACCCAGAGGGCATGTTTC 2793
Qy 894 IleValPheThrAlaIleAlaYrGlyAlaMetAlaIleGlyLysThrLeuValLeuAla 913
Db 2794 ATAGTTTTTACTGCAATTCATATGGAGCTATGCCATCGAAAAACGCTCGTTTTGGCT 2853
Qy 914 ProGluYrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuGluLysLys 933
Db 2854 CTTGANTATCCAAAGCCAAATCGGGGGCTGCACATCTGTTGCTTGTGGNAAAGAA 2913
Qy 934 ProAsnIleAspSerArgSerGlnGluGlyLysLeuProAspThrCysGluGlyAsnLeu 953
Db 2914 CCAAAATATAGACGCGCAGTCAAGAAGGCAAAAGCCAGACACATGTGAAGGGAATTTA 2973
Qy 954 GluPheArgGluValSerPhePheYrProCysArgProAspValPheIleLeuArgGly 973
Db 2974 GAGTTTCGAGAGTCTCTTTCTTCATCCATGTCGCCACAGATGTTTCATCTCCCTGGC 3033
Qy 974 LeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGly 993
Db 3034 TTATCCTCAGTATTGAGCGAGGAGAGACAGTAGCATTTGTGGGGAGACGCGCTGGG 3093
Qy 994 LysSerThrSerValGlnLeuLeuGlnArgLeuYrAspProValGlnGlyGlnValLeu 1013
Db 3094 AAAAGCACTTCTGTTCAACTTCTGCAGAGACTTTATGACCCCTGCAAGGACAAAGTGCTG 3153
Qy 1014 PheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIle 1033
Db 3154 TTTGATGGTGTGGATGCAAAAGAAATTTGAATGTACAGTGGCTCCGTTCCCAAAATAGCAATC 3213
Qy 1034 ValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaYrGlyAsp 1053
Db 3214 GTTCTCTAAGAGCCTGTGCTCTTCAACTGCAGCATTTGCTGAGAACATCGCCTATGTTGAC 3273
Qy 1054 AsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHis 1073
Db 3274 AACAGCGGTGTGTGCAATTAGATGAGTCAAGAAGCCGCAAAATGCAGCAATATCCAT 3333
Qy 1074 SerPheIleGluGlyLeuProGluLysYrAsnThrGlnValGlyLysGlyAlaGln 1093
Db 3334 TCTTTTATTGAAGGTCTCCCTGAGAAATAACAACACAAAGTTGGACTGGAAGGAGCACAG 3393
Qy 1094 LeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuGlnLysProLys 1113
Db 3394 CTTTCTGGCGGCGCAGAAACAAGACTAGTATTGCAAGGGCTCTTCTCCAAAAACCCAAA 3453
Qy 1114 IleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGln 1133
Db 3454 ATTTTATTGTTGAGGCCACTTCAGCCCTCGATATGACAGTGGAAGGTGTTTCAG 3513
Qy 1134 HisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSer 1153
Db 3514 CATGCCCTTGATAAAGCCAGGACGGGAAGACATGCTAGTGTGTCATCTCACAGGCTCTCT 3573

Qy 1154 AlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGly 1173
Db 3574 GCAATTTCAGAACGCAGATTGTAGTGGTCTTCGCAATGGAAGATAAAGGAACAAGGA 3633
Qy 1174 ThrHisGlnGluLeuLeuArgAsnArgAspIleYrPheLysLeuValAsnAlaGlnSer 1193
Db 3634 ACTCATCAAGAGCTCCTGAGAAATCGAGACATATATTTTAAGTTAGTGAATGCACAGTCA 3693
Qy 1194 ValGln 1195
Db 3694 GTGCAG 3699
RESULT 3
US-09-873-409-12
; Sequence 12, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
; APPLICANT: Frank, Markus
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873,409
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 3177
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: No. US20020037522A1e
; LOCATION: (198)..(198)
; OTHER INFORMATION: n at position 198 represents any nucleotide (A, T, C or G)
US-09-873-409-12
Alignment Scores:
Pred. No.: 0 Length: 3177
Score: 5159.50 Matches: 1031
Percent Similarity: 97.45% Conservative: 0
Best Local Similarity: 97.45% Mismatches: 0
Query Match: 85.82% Indels: 27
DB: 10 Gaps: 1
US-09-873-409-6 (1-1195) x US-09-873-409-12 (1-3177)
Qy 165 MetValIleSerLeuThrSerLysGluLeuSerAlaYrSerLysAlaGlyAlaValAla 184
Db 1 ATGGTCATCTCATTTGACCAGTAAGGAATTAAGTGCCTATTCCAAAGCTGGGCTGTGGCA 60
Qy 185 GluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeu 204
Db 61 GAAGAAGTCTTGTTCATCAATCCGACAGTCATAGCCCTTTAGGGCCCCAGGAGAAAGAACTT 120
Qy 205 Gln----- 205
Db 121 CAAAGTCTTCTCTTTAAATAACAAGATATGCTTGGTTTATTTTCCCCAGTGGCTA 180
Qy 206 -----ArgYrThrGlnAsnLeuLysAspAlaLysAspPhe 217
Db 181 CTAAGTTGTGTTCTGTTNTTTGTAAGGTATACACAGAACTCTCAAGATGCAAAAGGATTTT 240
Qy 218 GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValYrPhePheMetAsn 237
Db 241 GGCATAAAAGGACTATAGCTTCAAAAAGTGTCTCTGGGTGCTGTGTACTTCTTTATGAAT 300
Qy 238 GlyThrYrGlyLeuAlaPheTrpYrGlyThrSerLeuIleLeuAsnGlyGluProGly 257
Db 301 GGNACCTATGGACTTGTCTTTTGGTATGGAACCTCCTTCATCTTAATGGAGAACCTGGA 360
Qy 258 TyrThrIleGlyThrValLeuAlaValPheSerValIleHisSerSerYrCysIle 277

Db 361 TATACCATCGGAGCTGTCTCTGCTGTTTTCTTTAGTGTAAATCCATAGCAGTATTATGCATT 420
Qy 278 GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 297
Db 421 GGAGCAGCAGTCCCTCACTTTGAAACCTTCGCAATAGCCCGAGAGCTGCTTTCATATT 480
Qy 298 PheGlnValIleAspLysProSerIleAspAsnPheSerThrAlaGlyTyrLysPro 317
Db 481 TTCAGGTTATTGATAGAAACCCAGTATAGATAAATTTTTCCACAGCTGATATAAACCT 540
Qy 318 GluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgPro 337
Db 541 GAATCCATAGAGAACTGTGGAAATTTAAAAATGTTTTCTTCAATTATCCATCAAGACCA 600
Qy 338 SerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu 357
Db 601 TCATATCAAGATTTCTGAAGGCTCTGAATCTCAGAAATTAAGTCTGGAGAGACAGTCGCCCTTG 660
Qy 358 ValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAsp 377
Db 661 GTCGGTCTCAATGGCAGTGGGAAGAGTACGTTAGTCCAGCTTCTGCAGAGGTTTATATGAT 720
Qy 378 ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis 397
Db 721 CCGGATGATGGCTTTATCATGTGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCAT 780
Qy 398 TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSer 417
Db 781 TATCGAGACCATATTGGAGTGGTGTAGTCAAGAGCGCTGTTTTGTTGGGACCAACATCAGT 840
Qy 418 AsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArg 437
Db 841 AACATATCAAGTATGACGAGATGATGTACTGATGAAGAGATGGAGAGCAGCBAAG 900
Qy 438 GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly 457
Db 901 GAAGCAAAATCGTATGATTTTATCATGGAGTTCCTTAATAAAATTAATCATTTGCTAGGG 960
Qy 458 GluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu 477
Db 961 GAAAAAGAGCTCAATAGTGGAGGGCAGAAACAGAGGATCGCAATTTGCTCGTCCCTTA 1020
Qy 478 ValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 497
Db 1021 GTTCGAACCCCAAGATCTCTGATTTTAGATGAGGCTACGCTCGCCCTGGATTTCAGAAAGC 1080
Qy 498 LysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValVal 517
Db 1081 AAGTCAGCTGTTCAAGCTGCACTGGAGAGCGCAGCAAAAGGTCGGACTACAATCGTGGTA 1140
Qy 518 AlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMet 537
Db 1141 GCACCCGACTTTCTACTATTTCGAAGTGCAGATTGTGATGACCCCTAAAGGATGGAATG 1200
Qy 538 LeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeu 557
Db 1201 CTGGCGGAGAAAGGAGCAGATCTGTAATTAATGGCAAAAACGAGGTCTATATTATTACATT 1260
Qy 558 ValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThr 577
Db 1261 GTGATGTCACAGGATATTAATAAAGCTGATGACAGATGGAGTCAATGACATATTCTACT 1320
Qy 578 GluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle 597
Db 1321 GAAAGAAAGACCAACTCACTTCCTCTGCTGCTGTAAGAGCATCAAGTCAGACTTCATT 1380
Qy 598 AspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeu 617
Db 1381 GACAAAGCTGAGGAAATCCCAATCTTAAGAGATAGTCTTCTGAAAGTCTCTCTATTATTA 1440
Qy 618 LysIleLeuLysLeuAsnLysProGluTyrProPheValValLeuGlyThrLeuAlaSer 637
Db 1441 AAAATTTAAAGTTAAACAAGCTGAATGGCTTTTGTGGTTCCTGGGGACATTTGGCTTCT 1500

Qy 638 ValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMet 657
Db 1501 GTTCTAAATGGAACTGTTTCATCCAGTATTTCATCATCATCTTTGCAAAATTTATAACCATG 1560
Qy 658 PheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePhe 677
Db 1561 TTTGGAAATTAATGATAAAACCCACATTAAGACATGATGCAGAAATTTATTCCATGATATTC 1620
Qy 678 ValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArg 697
Db 1621 GTCATTTTGGGTGTTATTGCTTTGTTCAGTTATTTTCATCGAGGATTTATTTACGGCAGA 1680
Qy 698 AlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGln 717
Db 1681 GCAGGGGAAATTTAACGATGAGATTAAGACACTTGGCCCTTCAAGCCCATGTTATATCAG 1740
Qy 718 AspIleAlaTyrPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAla 737
Db 1741 GATATTGCCCTGGTTTGTATGAAAAGGAAAAACAGCACAGGAGGCTTGCAACAATATTAGCC 1800
Qy 738 IleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuLeuThrGlnAsn 757
Db 1801 ATAGATATAGCAAAATTCAGGAGCAACAGGTTCAGGATTTGGCGTCTTTAACACAAAT 1860
Qy 758 AlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyrGluMetThrPhe 777
Db 1861 GCACATTAACATGGGACTTTTCAGTTATCTTCTTTATATATGATGGAGATGACATTC 1920
Qy 778 LeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMet 797
Db 1921 CTGATTTCTGAGTATTGCTCCAGTACTTGGCGTGACAGGAATGATTTGAAACCGCAGCAATG 1980
Qy 798 ThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGlu 817
Db 1981 ACTGGATTTGCCAACAAAGATAAGCAAGAACTTAAGCATGCTGGAAAGATAGCAACTGAA 2040
Qy 818 AlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMet 837
Db 2041 GCTTTGGAGAAATATAGTACTATAGTGTCAATTAACAGGGGAAAAAGCCTTCGAGCAATG 2100
Qy 838 TyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGly 857
Db 2101 TATGAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCACAGATATTGGA 2160
Qy 858 SerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPhe 877
Db 2161 AGCTGTATGCAATTCAGCCATGCCCTTTATATATTTTGGCTATGACAGGGGTTTCGATTT 2220
Qy 878 GlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThr 897
Db 2221 GGAGCCTATTAAATTCAGCTGGACGAATGACCCACAGAGGCAATGTTTCATAGTTTACT 2280
Qy 898 AlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSer 917
Db 2281 GCATTTGCATATGGAGCTATGGCATCGGAAAAACGCTCGTTTGGCTCTCGTAATATTC 2340
Qy 918 LysAlaLysSerGlyAlaAlaHisIlePheAlaLeuLeuGluLysLysProAsnIleAsp 937
Db 2341 AAAGCCAAATCGGGGCTCGCATCTGTTTGGCTGTGTAAGAAAGAAACCAATATAGAC 2400
Qy 938 SerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGlu 957
Db 2401 AGCCGAGCTCAAGAGGAAAGAACCCAGACACATGTCAGGAGNAATTTAGAGTTTCAGAA 2460
Qy 958 ValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSer 977
Db 2461 GTCTCTTCTCTCATCATGTCGCCAGATGTTTTCATCTCGTGGCTTATCCCTCAGT 2520
Qy 978 IleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSer 997
Db 2521 ATTGACGAGGAAAGACAGTAGCATTTGTGGGAGCAGCGGCTGTGGGAAAGACATCTCT 2580


```
QY 998 ValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlnGlyClnValLeuPheAspGlyVal 1017
Db 2581 GTTCAACTTCTGCAGAGACTTTATGACCCCGTGCAGAGGACAAGTGTGTTGATGGTGC 2640
QY 1018 AspAlaLysGluLeuLeuValGlnTyrLeuArgSerGlnLeuAlaLeuValProGlnGlu 1037
Db 2641 GATGCAAAAGAATTGAATGTACAGTGGCTCCGTTCCCAATAGCAATCGTTCTCAAGAG 2700
QY 1038 ProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgVal 1057
Db 2701 CCTGTGCTCTCAACTGCAGCATTTGTCGAAACATCGCCTATGTCGACCAACAGCGGTGC 2760
QY 1058 ValProLeuAspGluLeuLysGluAlaAlaAsnAlaAsnIleHisSerPheIleGlu 1077
Db 2761 GTGCCATTAGATGAGATCAAGAAGCGCAAAATGCAGCAAAATATCCATTCTTTTATTGAA 2820
QY 1078 GlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGly 1097
Db 2821 GGTCTCCCTGAGAAATACAAACACAAAGTTGGACTGAAGGAGCACAGCTTCTCGCGGC 2880
QY 1098 GlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeu 1117
Db 2881 CAGAAACAAAGACTAGCTATTGCAAGGCTCTTCTCCAAAAACCCCAAAATTTTATTGTTG 2940
QY 1118 AspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAsp 1137
Db 2941 GATGAGGCCACTTTCAGCCCTCGATATGACAGTGAAGAGGTGTTTCAGCATGCCCTTGAT 3000
QY 1138 LysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsn 1157
Db 3001 AAGCCAGGACGGGAGGACATCCCTAGTGGTCACACAGGCTCTCTGCAATTCAGAAC 3060
QY 1158 AlaAspLeuIleValValLeuHisAsnGlyLysIleLysGlnGlnGlyThrHisGlnGlu 1177
Db 3061 GCAGATTGTAGTAGTGGTCTGCACATATGNAAGATAAAGNAACAAGNACTCATCAAGAG 3120
QY 1178 LeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 1195
Db 3121 CTCTCGAGAAATCGACACATATATTTTAAGTTAGTGAATGCACAGTCAGTGCAG 3174
```

RESULT 4
US-09-873-409-10
; Sequence 10, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
; APPLICANT: Frank, Markus
; APPLICANT: Sayegh, Mohamed
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873,409
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 10
; LENGTH: 2856
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-873-409-10

Alignment Scores:
Pred. No.: 0 Length: 2856
Score: 4561.50 Matches: 912
Percent Similarity: 97.96% Conservative: 0
Best Local Similarity: 97.96% Mismatches: 1
Query Match: 75.87% Indels: 18
DB: 10 Gaps: 1

US-09-873-409-6 (1-1195) x US-09-873-409-10 (1-2856)

```
QY 266 ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHis-PheGI 285
Db 114 GTTCTTCTTAGTGTAAATCCATAGCAGTTATTGCAATTGGAGCAGCAGTCCCTCATTTATT-- 171
```

```
QY 285 utrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysPr 305
Db 172 -----GATAAGAAACC 182
QY 305 oSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGI 325
Db 183 CAGTATAGATAACTTTTCCACAGCTGGATATAAACCTGAATCCATAGAGAACTGTGA 242
QY 325 uPheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLe 345
Db 243 ATTTAAANAATGTTTCTTCAATTTATCCATCAAGACCATCTATCAGATTCTGAAAGTCT 302
QY 345 uAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLy 365
Db 303 GAATCTCAGAAATAAGTCTGGAGAGACAGTCGCTTGGTCGGTCTCAATGGCAGTCGGA 362
QY 365 sSerThrValValGlnLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVa 385
Db 363 GAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCGGATGATGGCTTTATCATGT 422
QY 385 lAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVa 405
Db 423 GGATGAGATGACATCAGAGCTTTAAATGTCGGCATTTATCGAGACCATATTGGAGTGT 482
QY 405 lSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAs 425
Db 483 TAGTCAAGAGCCTGTTTGTTCGGACCAACCATCAGTAACAATATCAAGTATGGACGAGA 542
QY 425 pAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIl 445
Db 543 TGATGTGACTGATGAAGAGATGGAGAGACAGCAAGGGAAAGCAAAATGCGTATGATTTAT 602
QY 445 eMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGI 465
Db 603 CATGGAGTTTCTTAATAATTAATACATTGTTAGGGGAAAAGAGAGCTCAAAATGAGTGG 662
QY 465 yGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIl 485
Db 663 AGGGCAGAAACAGAGATCGCAATTGCTCGTCCCTTAGTTTCGAACCCCAAGATTCTGAT 722
QY 485 eLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaIle 505
Db 723 TTTAGATGAGGCTAGCTCTGCCCTGGATTCAGAAAGCAAGTCAGCTGTTCAGAGCTGCAC 782
QY 505 uGluLysAlaSerLysGlyArgThrThrIleValAlaAlaHisArgLeuSerThrIleAr 525
Db 783 GGAGAAGGCGAGCAAAAGGTCGGACTACATCGTGGTAGCACACCCGACTTTCTACTATTGC 842
QY 525 gSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAl 545
Db 843 AAGTGCAGATTTGATTGTGACCTTAAGGATGGAATGCTGGCGGAGAAAGGAGCACATGC 902
QY 545 aGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIleLysLy 565
Db 903 TGAACCTAATGGCAAAACGAGGTCTATATTATTACCTTGTGATGTCACAGGATATTAATA 962
QY 565 sAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuPr 585
Db 963 AGCTGATGAACAGATGGAGTCAATGACATATTTCTACTGAAGAAAGAACCCACTCATCTTC 1022
QY 585 oLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThrGI 605
Db 1023 TCTGCACCTCTGTGAAGAGCATCAAGTCAGACTTCAITTCAGAGGCTGAGGAATCCACCCA 1082
QY 605 nSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysPr 625
Db 1083 ATCTAAAGAGATAAGTCTTCTCCTCAAGTCTCTCTATTAATAAAATTTTAAAGTTAAACAGCC 1142
QY 625 oGluTyrProPheValIleGlyThrLeuAlaSerValLeuAsnGlyThrValHisPr 645
Db 1143 TGAATGCCTTTTGTGGTTCCTGGGACATTTGGCTTCTGTCTTAAATGGAACTGTTTCATCC 1202
```


645 oValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrTh 665
Db 1203 AGTATTTTCCATCATCTTTGCAGAAATTAATACCATGTTTGGAAATAATGATAAAACCAC 1262
Qy 665 rLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPh 695
Db 1263 ATTAAGCATGATGACAGAAATTTATTCATGATATTCGTTCATTTGGGTGTTATTTGCTT 1322
Qy 685 eValSerTyrPheMetGlnGlyLeuPheTyrGlyAraGlyGluIleLeuThrMetAr 705
Db 1323 TGTCAAGTTATTTTCATGAGGATTTATTTACGGCAGACGAGGGAATTTTACGATGAG 1382
Qy 705 gLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGluLys 725
Db 1383 ATTAAGACACTTGGCTTCAAGCCATGTTATATCAGGATATTGCCTGGTTTTCATGAAA 1442
Qy 725 sGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnI 745
Db 1443 GGAACACAGCACAGGAGCTTGACAAATATTAGCCATAGATATAGCAAAATTTCAAGG 1502
Qy 745 yAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerVa 765
Db 1503 AGCAACAGGTTCCAGGATTCGCGCTTAACACAAATGCAAACTAACATGGGACTTTTCAGT 1562
Qy 765 lIleIleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProVa 785
Db 1563 TATCATTTCTTTATATATGATGGAGATGACATTCCTGATTCGATTTGCTCCAGT 1622
Qy 785 lLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLys 805
Db 1623 ACTTGGCGTGACAGCAATGATTAAGACCGCAGCAATGACTGGATTTGCCAACAAAGATAA 1682
Qy 805 sGlnLeuLysHisIleGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrI 825
Db 1683 GCAAGAACTTAAGCATCTGGAAGATAGCAACTGAAAGCTTTGGAGAAATATACGTACTAT 1742
Qy 825 eValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGl 845
Db 1743 AGTGCTATTAAACAGGAGGAAAGCTTCGAGCAATGATATGAGAGATGCTTCAGACTCA 1802
Qy 845 nHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAl 865
Db 1803 ACACAGAAATACCTCGAAGAAAGCACAGATATTATGGAAGCTGTTATGCAATTCAGCCATGC 1862
Qy 865 aPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrIleLeuGlnAlaGl 885
Db 1863 CTTTATATATTTTCCCTATGCAGCAGGTTTCGATTTTGGAGCTTATTTAATTTCAAGCTGG 1922
Qy 885 yArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAl 905
Db 1923 ACGAATGACCCAGCAGGCGATGTTTCATAGTTTTCCTGCAATTCATATGGAGCTATGGC 1982
Qy 905 aIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHi 925
Db 1983 CATCGAAAAACGCTCGTTTTGGCTCCTCGAATATTCCAAAGCCAAATCGGGGGCTGCGCA 2042
Qy 925 sLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLy 945
Db 2043 TCTGTTTCCCTTGTGGAAAAAGAAACCAAAATATAGACGCGCAGTCAAGAAAGGAAAAA 2102
Qy 945 sProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysAr 965
Db 2103 GCCAGACACATGTGAGGGAATTTAGAGTTTCGAGAGTCTCTTCTTCTATCCATGCTCG 2162
Qy 965 gProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAl 985
Db 2163 CCCAGATGTTTTTCATCTCCGCTGCTTATCCCTCAGTATTGAGCGAGGAAAGACAGTAGC 2222
Qy 985 aPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuGlnArgLeuTy 1005
Db 2223 ATTTGTGGGAGACGCGCTGTGGAAAAAGCACCTTCTGTTCACCTTCTGCAGAGACTTTA 2282
Qy 1005 rAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGl 1025

Db 2283 TGACCCCGTCGCAAGGACAGTCTGTTGATGCTGGATGCAAAAGAAATTGAATGTACA 2342
Qy 1025 nTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIl 1045
Db 2343 GTGGCTCCGTTCCCAAAATAGCAATCGTTCTCAAGAGCCTGTGCTCTTCACTGAGCAT 2402
Qy 1045 eAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGl 1065
Db 2403 TGTGAGACATCGCTATGTTGTCACACAGCCGTTGGTGCCATTAGATGAGATCAAAGA 2462
Qy 1065 uAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnTh 1085
Db 2463 AGCGCAAAATGACGCAAAATATCCATTTTATGAAGTCTCCCTGAGAAAATACAAACAC 2522
Qy 1085 rGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnAraGlyLeuAlaIleAl 1105
Db 2523 ACAAGTTGCACTGAAAGGAGCACAGCTTTCTGGCGCCAGAAACAAAGACTAGCTATTGC 2582
Qy 1105 aArgAlaLeuLeuGlnLysProLysIleLeuLeuAspGluAlaThrSerAlaLeuAs 1125
Db 2583 AAGGCTCTTCTCCAAAACCCCAAAATTTTATTTGGATGAGGCCACTTCAGCCCTCGA 2642
Qy 1125 pAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCy 1145
Db 2643 TAATGACAGTGAGAAGTGGTTTCAGCATGCCCTTGATAAGCCAGGACGGAAGGACATG 2702
Qy 1145 sLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHi 1165
Db 2703 CCTAGTGGTCACTCAAGGCTCTCTGCAATTCAGAACGAGATTTGATAGTGTCTGCA 2762
Qy 1165 sAsnGlyLysIleLysGluGlnGlyThrHisGlnLeuLeuAsnArgAspIleTy 1185
Db 2763 CAATGGAAGATTAAGAACACAGGAACTCATCAAGAGCTCTCTGAGAAATCGAGACATA 2822
Qy 1185 rPheLysLeuValAsnAlaGlnSerValGln 1195
Db 2823 TTTTAAGTTAGTGAATGCACAGTCAGTCAGCAG 2853

RESULT 5

US-09-866-866A-1
; Sequence 1, Application US/09866866A
; Patent No. US20020102244A1
; GENERAL INFORMATION:
; APPLICANT: Sorrentino, John
; APPLICANT: Schuetz, Brian
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
; FILE REFERENCE: 1340-1-021CIP2
; CURRENT APPLICATION NUMBER: US/09/866,866A
; CURRENT FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 09/584,586
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: PCT/US99/11825
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: 60/086,988
; PRIOR FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 3860
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-866-866A-1

Alignment Scores:
Pred. No.: 7,94e-311 Length: 3860
Score: 3554.00 Matches: 682
Percent Similarity: 75.63% Conservative: 246
Best Local Similarity: 55.58% Mismatches: 263
Query Match: 59.12% Indels: 36
DB: 10 Gaps: 9

US-09-873-409-6 (1-1195) x US-09-866-866A-1 (1-3860)

Qy	341	IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu	360
Db	1225	ATCTTGAAAGCCCTGAACCTGAAGCTGGCAGATGGGCACGGTGGCCCTGGTGGAAAC	1284
Qy	361	AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyTyrAspProAsp	380
Db	1285	AGTGGCTGTGGGAAGAGACACACAGTCCAGCTGTATGACAGAGGCTCTATGACCCACAGAG	1344
Qy	381	GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyTyrArgAsp	400
Db	1345	GGGATGTGTGCTGTATGGACAGGATATTAGGACCATAAATGTAAGCTTTCTACGGGA	1404
Qy	401	HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle	420
Db	1405	ATCATTTGCTGTGTAGTCAGGAACCTGTATTGTTTGGCCACCACAGATGATGAACAACATT	1464
Qy	421	LysTyTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn	440
Db	1465	CGCTATGGCCGTGAAATGTACCATGGATGAGATTGAGAAGCTGTCCAGGAAGCCCAAT	1524
Qy	441	AlaTyTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyLeuGly	460
Db	1525	GCCTATGACCTTATCATGAACTGCCTCATAAATTTGACACCCCTGTTGGAGAGAGAGGG	1584
Qy	461	AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn	480
Db	1585	GCCCACTGAGTGGTGGGAGAGCAGAGAGATCGCCATTGCACTGCGCCCTGGTTCGCAAC	1644
Qy	481	ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla	500
Db	1645	CCCAAGATCTCTCTGTGATGAGGCCAGCTCAGCCTTGGACACAGAAGCGAAGCACTG	1704
Qy	501	ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg	520
Db	1705	GTTTCAGTGGCTCTGGTAAGCCAGAAAGGTCGGACCCACCATTTGTGATAGCTCATCGT	1764
Qy	521	LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu	540
Db	1765	TTGTCTACAGTTCTGTAATGTGACGTCATCGCTGGTGTTCGATGATGGAGTCATTTGTGGAG	1824
Qy	541	LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyTyrSerLeuValMetSer	560
Db	1825	AAAGGAATCATGATGAACTCATGAAGAAGAAAGGCATTACTTCAAACTTGTCACAATG	1884
Qy	561	Gln-----AspIleLysLysAlaAspGluGlnMetGluSerMetThr	574
Db	1885	CAGACACAGGAATGAGTTGAATTTAGAAATGCGCTGATGAATCCAAAAGTGAAATT	1944
Qy	575	TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal-----	589
Db	1945	GATCCCTTGGAAATGCTCTTCAATGATTCAAGATCCAGCTCTAATAAGAAAAGATCAACT	2004
Qy	590	---LysSerIleLys---SerAspPheIleAspLysAlaGluLysSerThrGlnSerLys	607
Db	2005	CGTAGGAGTGTCCGTGGATCACAAAGCCCAAGACAGAAAGCTTAGTACCAAAGAGAGCTCTG	2066
Qy	608	GluIleSerLeuProGluValSerLeuLeuValIleLeuLysLeuAsnLysProGluTrp	627
Db	2065	GATGAAGATATACCTCCAGTTTCTTTTGGAGATTATGAGACTAAATTTAACTGAATGG	2127
Qy	628	ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe	647
Db	2125	CTTATTATTGTTGTTGTTGTTTGTGCCATTATATAATGGAGGCTGTCAACACGACATTT	2185
Qy	648	SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu	666
Db	2185	GCAATAATATTTTCAAAGATATAGGGGTTTTTACAAAGAAATTTGATGATCCTCGAAACAAA	2244
Qy	667	LysHisAspAlaGluIleTyTyrSerMetIlePheValIleLeuGlyValIleCysPheVal	686
Db	2245	CGACAGAAATAGTAACTGTGTTTTCACATTGTTTCTAGCCCTTGGAAATATTTCTTTTAT	2304
Qy	687	SerTyTyrPheMetGlnGlyLeuPheTyTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu	706


```
Db 2305 ACATTTTCTCAAGGTTTACATTTGGCAAGCTGGAGAGATCTCTCAACCAAGCGGCTC 2364
      :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Qy 707 ArgHisLeuAlaPheLysAlaMetLeuTyGlnAspIleAlaTrpPheAspGluLysGlu 726
      :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 2365 CGATACATGTTTCCGATCTCTCAGACAGAGATGGTGGTTGATGACCCCTAAA 2424
      :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Qy 727 AsnSerThrGlyGlyLeuThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 746
      :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 2425 AACACCACTGGAGCATTTGACTTACAGGCTCGCCATGATGCTGCTCAAGTTAAAGGGGCT 2484
      :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Qy 747 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValile 766
      :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 2485 ATAGGTTCCAGGCTTGCTGTAATTTACCAGCAATATAGCAAAATCTTGGGACAGGAATAAT 2544
      :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Qy 767 IleSerPheIleTyGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 786
      :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 2545 ATATCCTTTCATCTATGTTGGCAACTAACATGTTACTTTAGCAATTTGACCCCATTT 2604
      :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Qy 787 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaLeuLysAspLysGln 806
      :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 2605 GCATACACAGAGTGTGTGAATGAAATGTTGCTGCGACACGACTGGAAGATAGAAA 2664
      :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Qy 807 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 826
      :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 2665 GAACTAGAAGTGTGGGAAGATCGCTACTGAAGCAATAGAAAACCTCCGAACCGTTGTT 2724
      :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Qy 827 SerLeuThrArgGluLysAlaPheGluGlnMetTyGluGluMetLeuGlnThrGlnHis 846
      :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 2725 TCTTTGACTCAGGACGAGAAAGTTTGAACATATGTTATGCTCAGAGTTTTCAGGTTACCATAC 2784
      :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Qy 847 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyraLapheSerHisAlaPhe 866
      :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 2785 AGAACTCTTTGAGAAAAGCACATCTTTGGAATTACATTTTCCTTCAACCCAGGCAATG 2844
      :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Qy 867 IleTyraPheAlaTyraAlaAlaGlyPheArgPheGlyAlaTyraIleuLeuAlaGlyArg 886
      :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 2845 ATGATATTTTCTATGCTGGATGTTTCGGTTTGGAGCCTTACTTGGTGCCACATAAACTC 2904
      :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Qy 887 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyraGlyAlaMetAlaIle 906
      :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 2905 ATGAGCTTTGAGGATGTTCTGTTAGTATTTTTCAGCTGTTGCTTGGTGCATGGCCGCTG 2964
      :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Qy 907 GlyLysThrLeuValLeuAlaProGluTyraSerLysAlaLysSerGlyAlaAlaHisLeu 926
      :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 2965 GGGCAAGTCAGTTCTTCTGCTGATGTCCTGCAACCAAAATATCAGCAGCCACATC 3024
      :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Qy 927 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 946
      :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 3025 ATCATGATCATTTGAAAAAACCCCTTTGATTGACAGCTACAGCAGGAAGCGCTTAATGCCG 3084
      :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Qy 947 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyraProCysArgPro 966
      :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 3085 AACACATTTGGAAGAAATGTACATTTTGGTGAAGTTGATTCACTATCCACCCGACCG 3144
      :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Qy 967 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 986
      :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 3145 GACATCCAGTGTCTCAGGACTCAGCTGGAGTGTGAAGAGGCCAGACGCTGGCTCTG 3204
      :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Qy 987 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyraAsp 1006
      :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 3205 GTGGGACGACAGTGGCTGTGGGAAGACACAGTGGTCCAGCTCTCTGGAGCGGTTCTACGAC 3264
      :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Qy 1007 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 1026
      :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 3265 CCGTTGGCAGGAAAGTGTCTTGTATGGCAAGAAATAAAGCGACTGCAATGTTTCAGTGG 3324
      :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Qy 1027 LeuArgSerGlnIleAlaIleValProGlnProValLeuPheAsnCysSerIleAla 1046
      :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 3325 CTCGAGCACACCTGGGCATCTGTTCCAGGAGGCCATCTCTGTTGACTCGAGCATTTGCT 3384
      :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Qy 1047 GluAsnIleAlaTyraGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 1066
      :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
```

```
Db 3385 GAGAACATTCCTTATGGAGACAACACAGCCGGTGGTGTCTACAGGAAGAGATCGTAGGGCA 3444
      :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Qy 1067 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyraAsnThrGln 1086
      :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 3445 GCAAGGAGGCCAACATACATGCTTTCATCGAGTCACTGCTTAATAAATATAGCACTAAA 3504
      :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Qy 1087 ValGlyLeuLysGlyValaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 1106
      :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 3505 GTAGGAGACAAAAGAACTCAGCTCTCTGCTGGCCGAGAAACAACGCAATGCCATAGCTCGT 3564
      :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Qy 1107 AlaLeuLeuGlnLysProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspAsn 1126
      :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 3565 GCCTTTGTAGACAGCTCATATTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACA 3624
      :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Qy 1127 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1146
      :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 3625 GAAAGTGAAGAAGTTGTCGAAGAGCCCTGGACAAAGCCAGAGAGCCGACCTGCATT 3684
      :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Qy 1147 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1166
      :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 3685 GTGATTGCTCACCGCTGCTCCACCATCCAGAAATGCAGACTTAATAGTGGTGTTCAGAAAT 3744
      :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Qy 1167 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyraPhe 1186
      :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 3745 GCAGAGTCAAGAGGATGCGCAGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTT 3804
      :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Qy 1187 LysLeuValAsnAlaGlnSer 1193
      :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 3805 TCATGCTCAGTGTCCAGGCT 3825
      :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

RESULT 6
US-10-044-671-1
; Sequence 1, Application US/10044671
; Patent No. US20020177147A1
; GENERAL INFORMATION:
; APPLICANT: Washington State University Research Foundation
; APPLICANT: Mealey, Katrina
; APPLICANT: Bentjen, Steven
; TITLE OF INVENTION: MDR1 VARIANTS AND METHODS FOR THEIR USE
; FILE REFERENCE: 4630-61733
; CURRENT APPLICATION NUMBER: US/10/044,671
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/261,578
; PRIOR FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: US 60/314,829
; PRIOR FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1
; LENGTH: 4317
; TYPE: DNA
; ORGANISM: Canis familiaris
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (70)..(3912)
; OTHER INFORMATION:
US-10-044-671-1

Alignment Scores:
Pred. No.: 2,65e-310 Length: 4317
Score: 3549.00 Matches: 680
Percent Similarity: 74.65% Conservative: 239
Best Local Similarity: 55.24% Mismatches: 270
Query Match: 59.03% Indels: 42
DB: 9 Gaps: 8

US-09-873-409-6 (1-1195) x US-10-044-671-1 (1-4317)
Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
      :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 217 ATGTTGGTGGGACATGGCTGCATCATCCATGAGCTGCATCTCCCTCTCATGATGCTG 276
      :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Qy 21 ValLeuGlyGluMetSerAsnLeuIleSerGlyCysLeu----- 34
```


Db 277 GTTTTGGAAACATGACATAGCTTTGCAAAATGCGAGGAATTTCAAGAAACAAAACCTTTT 336
Qy 35 -----ValGlnThrAsnThrTyrSerPhePheArg----- 44
Db 337 CCAGTTATTAATTAAGAAAGTATATAGCAACATATACACACATTTTCATCAACCATCTGGAG 396
Qy 45 -----LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe 59
Db 397 GAGGAAATGACCACCATGATCCCTATTATTACAGTGGGATCGGTGCTGCGCTGGTGGCT 456
Qy 60 GlyTyrIleGlnIleSerLeuTyrIleThrAlaAlaArgGlnThrIlysergileArg 79
Db 457 GCTTACATCCAGGTTTCATCTCGTGGCTGGCAGCAGGAAGACAGATCTCAAAATTAGA 516
Qy 80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIle 99
Db 517 AAACAATTTTTCATGCTATCATGCGACAGGAGATGGCTGTGTCAGCTGCATGACGTT 576
Qy 100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
Db 577 GGGAGCTTAACACCCGGCTCACAGACGATGCTCCAAAATCAATGAAGGAATTGGCGAC 636
Qy 119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
Db 637 AAAGTTGGAATGTTCTTCAATCAATAGCAACATTTTTCACCGGTTTATAGTGGGGTTT 696
Qy 139 ValLysGlyTyrPheLeuThrValLeuLeuSerThrSerProLeuIleMetAlaSer 158
Db 697 ACACCTGGTTGGAAGTAAACCTTGTGATTTGGCCATCAGCCCTGTTCTTGACCTTTCA 756
Qy 159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
Db 757 GCGCCATCTGGCAAGATACTATCTTCTATTCTGATAGAAACTCTTGGCCATGCA 816
Qy 179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198
Db 817 AAAGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGAACTGTGATGCTTGGGA 876
Qy 199 AlaGlnGluGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 218
Db 877 GGACAAAGAAAGAACTTGAAGGTACACAAAAATTTAGAAGAGCTAAAAGAAATTGGG 936
Qy 219 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPheMetAsnGly 238
Db 937 ATAAAGAAAGCTATCACGCCCAACATTTCTATTGGTGGCCCTTCTTATTGATCATGCA 996
Qy 239 ThrTyrGlyLeuAlaPheTyrPyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 258
Db 997 TCATATGCTCTGGCTTCTGGTATGGGACCTCTTGCTCTCTCCAGTGAA-----TAT 1050
Qy 259 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 278
Db 1051 ACTATTGGACAGTACTCACTGCTCTTCTTCTGTATTAATTTGGGGCTTTTAGTATGGA 1110
Qy 279 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 298
Db 1111 CAGGCATCCCAAGCATTGAGCATTGTGAAAGCGAAGAGGACGACCTTATGAAATCTTC 1170
Qy 299 GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 318
Db 1171 AAGATAATTGACAATAAACCAAGCATTGCAGCTATTCGAAGAGTGGACATAAACAGAT 1230
Qy 319 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 338
Db 1231 AATATTAAAGGAAATTTGGAATTTCAAAATATGTTTCACTTCAGTTACCTCTTCGAAAAGAA 1290
Qy 339 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 358
Db 1291 GTTAAGATCTTAAGGGTCTCAACCTGAAGTTCAGAGTGGCGAGACAGTGGCGCTGGTT 1350
Qy 359 GlyLeuAsnGlySerGlyLysSerThrValGlnLeuLeuGlnArgLeuTyrAspPro 378

Db 1351 GGGAAACAGTGGCTGGGGAAGACGACCGTGCAGCTGATCCAGAGGCTCTATGACCCC 1410
Qy 379 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 398
Db 1411 ACAGATGGCATGGTCTGTATTGATGGACAGGACATTAGGACCAATAATGTAAGGCATCTT 1470
Qy 399 ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn 418
Db 1471 CGGGAATTTACTGGTGTGTGAGTCAGGAGCCTGTGTGTTTGGCCACCATAGTAGCTGAA 1530
Qy 419 AsnIleLysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGlu 438
Db 1531 AACATTGCTATGGCGGCAAAATGTCCATGGATGAGATTGAGAAAGCTGTTAAGGAA 1590
Qy 439 AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 458
Db 1591 GCCAATGCTATGATTTTATCATGAACCTACCTAATAATTTGACACTCTGTTGGAGAG 1650
Qy 459 LysGlyAlaGlnMetSerGlyGlyGlnArgIleAlaIleAlaArgAlaLeuVal 478
Db 1651 AGAGGGCGCCGCTGAGTGGTGACAGAAACAGAGAATCGCCATTTGCTCGGCCCTGGTT 1710
Qy 479 ArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 498
Db 1711 CGCAACCCCAAGATTTCTGCTGGATGAGGCAAGCTCAGCTCTGGACACTGAAAGTGAA 1770
Qy 499 SerAlaValGlnAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAla 518
Db 1771 GCAGTGGTTCAGTGGCCCTGGATGAAGCCAGAAAGCGGACTACCATTTGTAGAGCT 1830
Qy 519 HisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeu 538
Db 1831 CATCGTTGCTACAGTTCGTAAATGCCGATGTCAITGCTGTTTGTGATGATGAGTCAIT 1890
Qy 539 AlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuVal 558
Db 1891 GTGGAGAAAGAAATCATGATCAACTCATGAAGAGAGGACATTTTACTTCAAACTGTC 1950
Qy 559 MetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu 578
Db 1951 ACAATGCCAG---ACAAGAGGAATGNAATTGATTAGAAAATGCCACTGGTGAATCCAA 2007
Qy 579 ArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle--- 597
Db 2008 AGTGAAAGTGATGCCTTGGAAATG---TCTCCAAAAGATTACGGGTCCAGTTTAAATAAA 2064
Qy 598 -----AspLysAlaGluGluSer 603
Db 2065 AGAAGATCAACTCGCAGGAGTATACATGCACCAAGGCCAACAGACAGAAAGCTTGTGATA 2124
Qy 604 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLysIleLeuLysLeuAsn 623
Db 2125 AAAGAGGACTTGAATGAGAAATGTTCTTTCAGTTTCTTCTGGAGGATTTCTGAAGCTGAAC 2184
Qy 624 LysProGluTyrProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 643
Db 2185 TCAACTGAATGGCTATTATTTGTTGGTGTATATTTTGTGCTATTATATAAACGAGGCTG 2244
Qy 644 HisProValPheSerIleIlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAsp 662
Db 2245 CAACACGACATTTTCAATAATATTTTCAAGGATATAGGATCTTTACCCGAGATAGGAT 2304
Qy 663 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 682
Db 2305 CCTGAAACAAACGACAGAAATGATTAACATGTTTTCGTATTGTTTCTAGTCTTGGAAAT 2364
Qy 683 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 702
Db 2365 ATTTCTTTTATTACATTTTCTCCAGGGCTTCACATTTGGCAAGCTGGGGAGATCCTC 2424
Qy 703 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPhe 722
Db 2425 ACTAAGCGGCTTCGATACATGTTTTCAGATCCATGCTGAGACAGGATGTCAGCTGGTTT 2484


```
QY 723 AspGluysGluAsnSerThrGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 742
Db 2485 GATGACCTAAACAAACACCACCTGGAGCATTGACAAACAGGCTTCCCAATGATGGCTCAA 2544
QY 743 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 762
Db 2545 GTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTCATTACCCAGAGATATAGCAAAATCTGGG 2604
QY 763 LeuSerValIleIleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIle 782
Db 2605 ACAGGCATTATATATCTTAATCTATGTTGGCAATTAACACTTTTACTCTTAGCAAT 2664
QY 783 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 802
Db 2665 GTACCCATCATTCGATAGCAGAGTGTGTAATGAAATGTTGCTGGCAAGCACTG 2724
QY 803 LysAspLysGlnGluLeuLysHisAlaGlyIleAlaThrGluAlaLeuGluAsnIle 822
Db 2725 AAAGATAAGAAAGAGCTAGAAAGGAGCTGGGAAGATTGCTACAGAAGCCATCGAAACTTC 2784
QY 823 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluMetLeu 842
Db 2785 CGAACTGTTCTTGTGACTCGGGAGCAGAAAGTTGAATACATGATGACAGAGTTG 2844
QY 843 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 862
Db 2845 CAAGTACCATACAGAACTCTTTGAGGAAGACACACATCTTCGGGGTCTCATTTCTATC 2904
QY 863 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 882
Db 2905 ACCCAGCAATGATGATATTTTCTATGCTGGCTGTGTTCCGGTGTGCTACTTGGTG 2964
QY 883 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 902
Db 2965 GCAATGAGTTCATGAACCTTTCAGGATGTTCTTTGGTATCTCAGCTATGCTTTGGT 3024
QY 903 AlaMetAlaIleGlyLysLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 922
Db 3025 GCCATGCGCAGTGGGCGAGGTCAGTTTCATTTGCTCCTGACTATGCCAAGCCAAAGTATCA 3084
QY 923 AlaAlaHisIlePheAlaLeuLeuGluLysIleProAsnIleAspSerArgSerGlnGlu 942
Db 3085 GCAGCCACGCTCATCATGATGATGAAAAAGCCCTCTGATTGACAGCTACAGCCCTCAC 3144
QY 943 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPheTyr 962
Db 3145 GGCTCAAGCCAAATACGTTGGGAAGGAATGTGACATTTAATGAGGTCGTGTTCAACTAT 3204
QY 963 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 982
Db 3205 CCNACTCGACGACGATCCCGTGTCTCCAGGGCTGAGCTCGAGGTGAAGAGGGCCAG 3264
QY 983 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuGln 1002
Db 3265 ACGCTGGCCCTCGTAGTAGCAGTGGCTGTGGGAAGAGCAGATTGTCAGCTCCTAGAG 3324
QY 1003 ArgLeuTyrAspProValGlnGlnValLeuPheAspGlyValAspAlaLysGluLeu 1022
Db 3325 CGCTTCATGACCCCTTGGCTGGTTCAGTGCTAATGATGCCAAGAGATAAAGCACCTG 3384
QY 1023 AsnValGlnTyrLeuArgSerGlnAlaIleValProGlnGluProValLeuPheAsn 1042
Db 3385 AATGTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCTCAGGAGCCCATCTCTGTGAC 3444
QY 1043 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 1062
Db 3445 TGCAGCATTCGCGAACAATTCCTATGAGACAAACAGCCGGGTCTGATCACATGAAGAG 3504
QY 1063 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 1082
Db 3505 ATTATGAGGCGAGCCAAAGGAGCCCAACATACACCCTTCATCGAGACACTCCCTCGAGAA 3564
```

1083 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 1102
Db 3565 TACAACACCAGTAGTAGGACAAAGAAACCCAGCTCTCTGGTGGCCAGAAACAGCGCAT 3624
QY 1103 AlaIleAlaArgAlaLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 1122
Db 3625 GCATAGCTCGGCGCTCTTGTGTAGACAGCCCTCATATTTGCTTTGGATGAGCTACATCA 3684
QY 1123 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 1142
Db 3685 GCTCTGATACAGAAAGTGAAAGGTTGTCCAAGAGCCCTGGCAACAGCCAGAGAAGGC 3744
QY 1143 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 1162
Db 3745 CGCACCTGATTTGATCGCCCGCTTGTCCACCATCCAGATGCAGATTTAATAGTG 3804
QY 1163 ValLeuHisAsnGlyLysIleLysGlnGlnGlyThrHisGlnGlnLeuLeuArgAsnArg 1182
Db 3805 GTGTTTCAGAAATGGCAAGTCAAGGAGCATGGCACATCAACAGCTGCTGGCTCAGAAA 3864
QY 1183 AspIleTyrPheLysLeuValAsnAlaGlnSer 1193
Db 3865 GGCATCTATTTTCCATGATCAGTGTCCAGGCT 3897

RESULT 7
US-09-866-866A-3
; Sequence 3, Application US/09866866A
; Patent No. US20020102244A1
; GENERAL INFORMATION:
; APPLICANT: Sorrentino, Brian
; APPLICANT: Schuetz, John
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
; FILE REFERENCE: 1340-1-021CIP2
; CURRENT APPLICATION NUMBER: US/09/866, 866A
; CURRENT FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 09/584, 586
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: PCT/US99/11825
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: 60/086, 988
; PRIOR FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 3860
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-866-866A-3

Alignment Scores:
Pred. No.: 5,16e-310 Length: 3860
Score: 3545.00 Matches: 681
Percent Similarity: 75.55% Conservative: 246
Best Local Similarity: 55.50% Mismatches: 264
Query Match: 58.97% Indels: 36
DB: 10 Gaps: 9

US-09-873-409-6 (1-1195) x US-09-866-866A-3 (1-3860)

QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 151 ATGGTGGTGGGAACCTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 210
QY 21 ValLeuGlyGluMetSerAsp 31
Db 211 GTGTTTGGAGAAATGACAGATATCTTTGCAATGACAGAAATTTAGAGATCTGATGTCA 270
QY 32 GlyCysLeuValGlnThr 43
Db 271 AACATCACTAATAGAACTGATATCAATGATACAGGGTCTTCTCATGAATCTGGAGGAAGAC 330
QY 44 -----ArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db 44 -----ArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61

Db 331 ATGACCCAGATATGCTTATTATTACAGTGGAAATTGGTGTGGGGTGTGCTGCTGCTTAC 390
Qy 62 ILeGlnIleSerLeuTriPleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db 391 ATTCAGGTTTCATTTTGGTCTGGCAGCTGGAACAAATACACAAAATTAGAAAACAG 450
Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db 451 TTTTTCATGCTATTAATGCGACAGAGATAGGCTGCTGTTGTGTCAGCATGTTGGGAG 510
Qy 102 LeuAenThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db 511 CTTAACCCCGACTTACAGATGATGCTCTTAAGATTAAATGAAGTTATTGGTGACAAAATT 570
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValIys 140
Db 571 GGAATGTTCTTTCAGTCAATGCAACATTTTTCACCTGGGTTTATAGTAGGATTACAGCT 630
Qy 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAla 160
Db 631 GGTGGAGCTAAACCTTGTGATTTGGCCATCAGTCTCTGTTCTTGGACTGTGCTGCT 690
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaThrSerLysAla 180
Db 691 GTCTGGGCAAGATACTATCTTCATTTACTGATAAAGAACTCTTACGCGTATCAAAAGCT 750
Qy 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db 751 GGAGCAGTAGCTGAAGAGGCTCTGGCAGCAATTAGACCTGTGATGTCATTTGGAGACAA 810
Qy 201 GluLysGluLeuGlnArgPyrThrGlnAenLeuLysAspAlaLysAspPheGlyIleLys 220
Db 811 AAGAAGAACTTTGAAAGGTACAAACAAATTTAGAAAGAGCTTAAAGAAATTTGGGATAAG 870
Qy 221 ArgThrIleAlaSerLysValSerLeuGlyAlaValThrPhePheMetAenGlyThrTyr 240
Db 871 AAAGCTATTACAGCCAAATATTTCTATAGTGTGCTGCTTCTCTGATCATCATCTTAT 930
Qy 241 GlyLeuAlaPheTrpPyrGlyThrSerLeuIleLeuAenGlyGluProGlyThrIle 260
Db 931 GCTCTGGCCCTCTGGTATGGGACCACCTTGGTCTCTCAGGGAA-----TATTCATT 984
Qy 261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
Db 985 GCACAAGTACTCACTGATCTTCTTCTGATTAAATTTGGGGCTTTTAGTCTGCGACGGCA 1044
Qy 281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
Db 1045 TCTCCAAGCATTTGAAGCATTTGCAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1104
Qy 301 IleAspLysPheProSerIleAspAenPheSerThrAlaGlyTrpLysProGluSerIle 320
Db 1105 ATTGATAATAAGCAAGTATTGACACTATTGCAAGGTGGGACCAACCAACAGATAATATT 1164
Qy 321 GlyGlyThrValGluPheLysAenValSerPheAenTyrProSerArgProSerIleLys 340
Db 1165 AAGGCAAAATTTGGAATTCAGAAATGTTCACTTCAGTACCCTCTCGAAAGAAAGTTAAG 1224
Qy 341 IleLeuLysGlyLeuAenLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
Db 1225 ATCTGAAGGGCTGAACCTGAAGGTGCGAGGTGGGACAGCGTGGCCCTGGTTGGAAC 1284
Qy 361 AenGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
Db 1285 AGTGGCTGTGGGAAGACCAACAGCTCCAGCTGATGACAGGGCTCTATGACCCCAACAGAG 1344
Qy 381 GlyPheIleMetValAspGluAenAspIleArgAlaLeuAenValArgHisTyrArgAsp 400
Db 1345 GGGATGTCAGTGTTCATGGACAGGATATTAGACCATAAATGTAAGTTTCTCTCGGAA 1404
Qy 401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAenAsnIle 420
Db 1405 ATCATTTGGTGTGGTGTGAGTCAGGAACCTGATTGTTGCGACACAGCATAGTGAACATTT 1464

Qy 421 LysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAen 440
Db 1465 CGCTATGCGCGTGAATAATGTCAACATGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAAT 1524
Qy 441 AlaTyrAspPheIleMetGluPheProAenLysPheAenThrLeuValGlyGluLysGly 460
Db 1525 GCCTATGACTTTTATCATGAACTGCTCATTAATTTGACCCCTGTTGGAGAGAGAGGG 1584
Qy 461 AlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAen 480
Db 1585 GCCCAGTTGAGTGTGGGACAGAGGATCGCAATTCACAGTGCCTGTGTTGCGAAC 1644
Qy 481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db 1645 CCCAAGATCTCTGCTGATGAGGCCACGTCAAGCTTTGGACACAGAAAGCGAAGCAGTG 1704
Qy 501 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 520
Db 1705 GTTCAGGTGGCTCTGATAAGCCCAAGAAAGGTGCGACCCCATTTGTATAGTCATCGT 1764
Qy 521 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 540
Db 1765 TTGCTACAGTTCGTAATGCTGACGTCTCGCTGCTTCGATGATGGAGTCATTTGGAG 1824
Qy 541 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerIleValMetSer 560
Db 1825 AAAGGAAATCATGATGAATCATGAAAGAGAAAGGCAATTTACTTCAAACTTGTCACAATG 1884
Qy 561 Gln-----AspIleLysLysAlaAspGluGlnMetGluSerMetThr 574
Db 1885 CAGACAGCAGGAAATGAAGTTGAATTTAGAAATGACGCTGATGAAATTCACAAAGTGAAT 1944
Qy 575 TyrSerThrGluArgLysThrAenSerLeuProLeuHisSerVal----- 589
Db 1945 GATGCTTGGAAATGCTTCAATGATTCAGATCCAGTCTTAATAAGAAAAGATCAACT 2004
Qy 590 ---LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys 607
Db 2005 CGTAGGAGTGTCCGTGGATCAACAGCCCAAGACAGAAAGCTTAGTACCAAGAGGCTCTG 2064
Qy 608 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAenLysProGluTrp 627
Db 2065 GATGAAAGTATACCTCCAGTTTCTTTTGGAGGATTTAGAGCTAAATTTAACTGAATGG 2124
Qy 628 ProPheValValLeuGlyThrLeuAlaSerValLeuAenGlyThrValHisProValPhe 647
Db 2125 CTTTATTTTGTGTTGGTGTATTTGTGCCATTATAATGGAGGCTGCAACAGCATTT 2184
Qy 648 SerIleIlePheAlaLysIleIleThrMetPheGlyAen---AsnAspLysThrThrLeu 666
Db 2185 GCATAATATTTTCAAGATTTATAGGGGTTTTTACAAGAATTTGATGATCCTGAAACAAA 2244
Qy 667 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 686
Db 2245 CGACAGAAATAGTAACCTGTTTTCACCTATTGTTTCTAGCCCTTTGGAATTTATTTTAT 2304
Qy 687 SerTyrPheMetGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 706
Db 2305 ACATTTTCTTCAAGTTTTCATTTTGGCAAGCTTGGAGAGATCTTCAACAGCGGCTC 2364
Qy 707 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 726
Db 2365 CGATACATGTTTTCGATCCATGCTCAGACAGGATGTGAGTTGGTTGATGACCCTAAA 2424
Qy 727 AenSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 746
Db 2425 AACACCATCTGGAGCATTCGACTACAGGCTCGCCAAATGATGCTGCTCAAGTTAAAGGGCT 2484
Qy 747 ThrGlySerArgIleGlyValLeuThrGlnAenAlaThrAenMetGlyLeuSerValIle 766
Db 2485 ATAGGTTCCAGGCTTGTGTAATTTACCCAGAAATATAGCAAAATCTTTGGGACAGGAATAT 2544

Qy	767	IleSerPheIleTyrGlyTyrGluMetThrPheLeuLeuLeuSerIleAlaProValLeu	786
Db	2545	ATATCTCTTCATCTATGGTTGGCAACTAAACACACTGTTACTCTTAGCAATTTGTATCCCATCAT	2604
Qy	787	AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln	806
Db	2605	GCAATAGCAGAGTGTGTGAATAGAAATGTTGTCTGGCAAGCACTGAAAGATAGAGAA	2664
Qy	807	GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal	826
Db	2665	GAACTAGAGGTGCTGGAGAGATCGCTACTGAAGCAATAGAAAACCTCCGGAACCGTTGTT	2724
Qy	827	SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis	846
Db	2725	TCTTTGACTCAGGAGCAGAAAGTTTGAACATATGTATGCTCAGAGTTTCCAGGTACCATAC	2784
Qy	847	ArgAsnThrSerLysLysAlaGlnIleLeuGlySerCysTyrAlaPheSerHisAlaPhe	866
Db	2785	AGAAACTCTTTGAGGAAGCACACATCTTTGGAAATACATTTTCTTCACCAGCGCAATG	2844
Qy	867	IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg	886
Db	2845	ATGTATTTTTCCTATGCTGATGTTTCCGGTTTGGNGCTACTTGGTGGCACAATAACTC	2904
Qy	887	MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle	906
Db	2905	ATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTTCTTTGGTGGCATGCCGCT	2964
Qy	907	GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu	926
Db	2965	GGCAAGTCAGTTCAATTTGCTCTGACTATGCCAAGCCAAATATATCAGCAGCCCAATC	3024
Qy	927	PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro	946
Db	3025	ATCATGATCATTTGAAAACCCCTTTGATTTAGAGCTACAGCAGCGAAGGCGCTAATCGCG	3084
Qy	947	AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro	966
Db	3085	AACACATTTGGAAGAAATGTACATTTTGGTGAAGTTGTATTCAACTATCCCAACCGACCG	3144
Qy	967	AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe	986
Db	3145	GACATCCAGTGCTTCAGGACATGACGCTTGGAGGTGGAAGGGCCAGACGCTGGCTCTG	3204
Qy	987	ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp	1006
Db	3205	GTGGGACAGCTGGCTGTGGGAAGACACACAGTGTCTCAGCTCTTGGACGGTTCTACGAC	3264
Qy	1007	ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTyr	1026
Db	3265	CCCTTTGGCAGGAAAGTGCTGCTTGATGTCAAAGAAATAAAGCGACTCAATGTTTCAGTGG	3324
Qy	1027	LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla	1046
Db	3325	CTCCGACACACCTGGGCATCTGTTGCCAGAGCCCATCTCTTTTGACTCGACGATTTGCT	3384
Qy	1047	GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla	1066
Db	3385	GAGAACATTTCCTATGGAGACACACCCGGGTGGTGTACAGGAAGAGATCTGAGGGCA	3444
Qy	1067	AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln	1086
Db	3445	GCAAAGGAGGCCACATACATGCTTTCATCGAGTCACTGCCTTAATAAATATAGCACTAAA	3504
Qy	1087	ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg	1106
Db	3505	GTAGGAGACAAAGAACTCAGCTCTCTGGTGGGCCAGAAACACGCAATTTGCCATAGCTCGT	3564
Qy	1107	AlaLeuLeuGlnLysProLysIleLeuLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn	1126
Db	3565	GCCCTTGTAGACAGCCTCATATTTTGTCTTTTGGATGAAGCCACGCTGAGCTCTGGATACA	3624
Qy	1127	AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu	1146

```

Db 3625 GAAAGTGAAGAGGTGCTCCAGAAAGCCCTGGGACAAAGCCAGAGAAAGCCGACCTGCATT 3684
Qy 1147 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1166
Db 3685 GTGATGCTCACCCTGTCACCACTCCGAATCGAGACTTAAATAGTGGTGTTCAGAAT 3744
Qy 1167 GlyIleLeuLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1186
Db 3745 GGCAGAGTCAAGGAGCATGGCAGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTT 3804
Qy 1187 LysLeuValAsnAlaGlnSer 1193
Db 3805 TCAATGTCAGTGTCCAGGCT 3825

RESULT 8
US-09-306-417-1
; Sequence 1, Application US/09306417
; Patent No. US20020103144A1
; GENERAL INFORMATION:
; APPLICANT: Heinrich-Pette-Institut
; TITLE OF INVENTION: Retroviral Gene Transfer Vectors
; FILE REFERENCE: P50491
; CURRENT APPLICATION NUMBER: US/09/306.417
; CURRENT FILING DATE: 1999-05-06
; EARLIER APPLICATION NUMBER: DE 198 22 115
; EARLIER FILING DATE: 1998-05-08
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 8630
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: proviral
; OTHER INFORMATION: plasmid DNA
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(160)
; OTHER INFORMATION: plasmid backbone (pUC)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (161)..(677)
; OTHER INFORMATION: 5'-LTR
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (532)..(1219)
; FEATURE:
; NAME/KEY: mat peptide
; LOCATION: (1220)..(5062)
; OTHER INFORMATION: m4 mdr-1 cDNA
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (5215)..(5774)
; OTHER INFORMATION: 3'-LTR
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (5775)..(8630)
; OTHER INFORMATION: plasmid backbone (pUC)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(8630)
; OTHER INFORMATION: retroviral expression vector SPbeta71m4
US-09-306-417-1

```


US-09-873-409-6 (1-1195) x US-09-306-417-1 (1-8630)

QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyValaCysLeuProLeuMetProLeu 20
DB 1370 ATGGTGGTGGGAACCTTGGCTGCCATCATCCATGGGGCTGGAGCTTCTCATGATGCTG 1429
QY 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
DB 1430 GTGTTGGAGAAATGACAGATATCTTTGCAAAATGCAGGAAATTTAGAAAGATCTGATGTCA 1489
QY 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
DB 1490 AACATCACTAATAGAGAGATATCATGATGATACAGGCTTCTTCATGAAATCTGGAGGAAGAC 1549
QY 44 -----ArgLeuThrLeuTyrTyrValGlyIleGlyValaAlaLeuIlePheGlyTyr 61
DB 1550 ATGACCAGATATGCCATTTATACAGTGGAAATGGTCTGGGGTGTGCTTGTCTGTAC 1609
QY 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
DB 1610 ATTCAAGTTTCATTTTGGTGGCTGGCAGCTGGGAAGACAAATACACAAATTTAGAAACAG 1669
QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
DB 1670 TTTTTCATGCTATATAGCGACAGGAGATAGGCTGGTTTGTATGTGCACGATGTGGGGAG 1729
QY 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
DB 1730 CTTAACCCCGACTTACAGATGATGCTCTTAAGATTAAATGAAGATTATGTGTACAAATTT 1789
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
DB 1790 GGAATGCTTTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGGATTTACACGT 1849
QY 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
DB 1850 GGTGGGAAGCTAACCTTGTGATTTTGGCAGTACAGTCTCTTGGTGGTGGTGGTGGTGGT 1909
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
DB 1910 GTCTGGGCAAGATACTATCTTCTTACTGTATAAGAACTCTTACGGTATGCAAAAGCT 1969
QY 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
DB 1970 GGAGCAGTAGCTGAAGAGCTTTGGCAGCAATTTAGAACTGTGATTCATTTGGAGACAA 2029
QY 201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
DB 2030 AAGAAGAACTTGAAGAGGTACACAAATAATTTAGAAAGAGCTTAAAGAAATTTGGGATAAG 2089
QY 221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
DB 2090 AAAGCTATTACAGCCAAATTTCTATAGTGTGCTTCTCTGCTGATCATCATCTTAT 2149
QY 241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
DB 2150 GCTGTCGCTCTCTGCTATGGGACCACTTGGTCTCTCAGGGGAA-----TATTCTATT 2203
QY 261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
DB 2204 GGACAAAGTACTCACTGATTTCTTTCTGTTAAATTTGGGGCTTTTATGTGTGGACAGGCA 2263
QY 281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
DB 2264 TCTCCAAGCATTTGAAGCATTTTGAATTTGAAGAGGAGCAGCTTATGAAATCTTCAAGATA 2323
QY 301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
DB 2324 ATTGATAATAAGCAAGTATTGACACTATTGCAAGAGTGGGCACAAACACAGATATATT 2383
QY 321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
DB 2384 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCACTTCACTTCACTTCACTTCACTTCACTT 2443

QY 341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
DB 2444 ATCTTGAAGGGCTGAACTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAAC 2503
QY 361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
DB 2504 ACTGCTGTGGGAAGAGCACACACAGTGCAGCTGCAGAGGCTCTATGACCCCAAGAG 2563
QY 381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
DB 2564 GGGATGGTCAAGTGTGATGCGACAGATATTAGGACCATAAATGTAAAGTTTCTACGGAA 2623
QY 401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnIle 420
DB 2624 ATCAITGGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2683
QY 421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
DB 2684 CCCTATGCCCGTGAATAATGTCCACATGATGAGATTTGAAAGCTGTCAAGGAAGCCAAT 2743
QY 441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGly 460
DB 2744 GCCTATGACTTTTATCATGAACTGCTCATAAATTTGACACCTCTGGTTGGAGAGAGAG 2803
QY 461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
DB 2804 GCCCAGTTGAGTGGTGGCAGAGCAGAGGATCGCATTTGCAGCTGGCTGGTTGCGCAAC 2863
QY 481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
DB 2864 CCAAGATCTCTCTGCTGGATGAGCCACGTCAGCTGAGTGGACACAGAAAGCGAAGCTG 2923
QY 501 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValAlaAlaHisArg 520
DB 2924 GTTCAGTGGCTCTCGATAAGGCCAGAAAGTGGACCCACCATTTGTATGATCATCT 2983
QY 521 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 540
DB 2984 TTGTCTACAGTTTCGTAATGCTGACGTCATCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3043
QY 541 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 560
DB 3044 AAAGAAATCATGATGAACCTCATGAAGAGAAAGGATTTACTTCAAACTTGTCAAAATG 3103
QY 561 Gln-----AspIleLysLysAlaAspGluGlnMetGluSerMetThr 574
DB 3104 CAGACAGCAGGAAATGAAGTTGAATTTAGAAATGCGAGCTGATGAATCCAAAGTGAAT 3163
QY 575 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal----- 589
DB 3164 GATGCTTGGAAATGCTTCAATGATTCACAGATCCAGTCTAATAAGAAAGAACACT 3223
QY 590 ---LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys 607
DB 3224 CGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACCAAGAGGCTG 3283
QY 608 GluIleSerLeuProGluValSerLeuLysIleLeuLysLeuAsnLysProGluTrp 627
DB 3284 GATGAAGTATACCTCCAGTTCCTTTTGGAGGATTTATGAAGCTAATAATTAATCTGAATGG 3343
QY 628 ProPheValValGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 647
DB 3344 CTTATTGTTGTTGGTGTATTGTCATATAAATGGAGGCTGCAACAGCATTT 3403
QY 648 SerIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 666
DB 3404 GCAATAATATTTTCAAGATTTATAGGGTGTTTTACAAGAAATTTGATCTCTGAAACAAA 3463
QY 667 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 686
DB 3464 CGACAGAAATAGTAACCTGTTTCACTATTGTTTCTAGCCCTTGAATATTATTTCTTTATT 3523


```
QY 687 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluLeuLeuThrMetArgLeu 706
Db 3524 ACATTTTCTTCAAGTTTTCACATTTGGCAAGCTGGAGAGATCTCCACCAAGCGGCTC 3583
QY 707 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 726
Db 3584 CGATACATGGTTTCCGATCCATCTCAGACAGAGTGTGAGTTGGTTGATGACCCCTAA 3643
QY 727 AspSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 746
Db 3644 AACACACTGGAGCATGACTACAGGCTCGCAATGATGCTCTCAAGTTAAAGGGGCT 3703
QY 747 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValile 766
Db 3704 ATAGGTTCCAGGCTTGTGTAATATCCAGCAATATAGCAAAATCTGGGACAGGAATAAT 3763
QY 767 IleSerPheIleTyrGlyTrpGluMetThrPheIleLeuLeuSerIleAlaProValLeu 786
Db 3764 ATATCTTCTATCTATGTTGGCAACTAACACTGTTACTCTTAGCAATTGTAACCATCAT 3823
QY 787 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 806
Db 3824 GCATAGCAGGAGTGTGTAATATGAATAATGTTCTGCGACAGCACTGAAGAATAGAAA 3883
QY 807 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 826
Db 3884 GAACTAGAGGTTCTGGGAAGATCGCTACTCTGAAGSAATAGAAAATCTCCGACCGTTGT 3943
QY 827 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluMetLeuGlnThrGlnHis 846
Db 3944 TCTTTGACTCAGGAGCAGAAATTTGAACATATATGATATGCTCAGAGTTTGCAGGTACC 4003
QY 847 ArgAsnThrSerLysLysAlaGlnIleLleGlySerCysTyrAlaPheSerHisAlaPhe 866
Db 4004 AGAACTCTTTGAGGAAGAACACACATCTTTGGAAATATACATTTTCCTCCAGGCAATG 4063
QY 867 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 886
Db 4064 ATGTAATTTTCTATGCTGATGTTTCCGGTTTGGAGCTTACTTGTGGGCACATAAATC 4123
QY 887 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 906
Db 4124 ATGAGCTTTGAGGATGTTCTGTAGTATTTTACGTGTTCTGTTGTCGTCATGGCGGTG 4183
QY 907 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 926
Db 4184 GGGCAAGTCAGTTCATTTGCTCTCTGATGATGCCAAGCCAAATATCAGCAGGCCACATC 4243
QY 927 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 946
Db 4244 ATCATGATCATTTGAAAACCCCTTTTGATTTGACAGCTACAGCAGGAAGCCCTAATGCCG 4303
QY 947 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 966
Db 4304 AACACATTGGAAGGAATGTACATTTTGGTGAAGTGTATTTCAACTATCCACCCGACCG 4363
QY 967 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 986
Db 4364 GATCCCATGCTCTCAGGAGCTGAGGCTGGAGGTGAAGAAGGGCCAGAGCTGGCTCTG 4423
QY 987 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 1006
Db 4424 GTGGCAGCAGTGGCTGTGGGAAGAGACACAGTGTCTCAGCTCTCTGGAGCGGTTCTACGAC 4483
QY 1007 ProValGlnGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 1026
Db 4484 CCCTTGGCAGGGAAGTGTCTGTATGTGTCGCAAGAAATTAAGCGACTGAATGTTCACTGG 4543
QY 1027 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 1046
Db 4544 CTCGAGCACACCTGGGCATCGTGTCCAGGAGCCCATCTGTTTGACTGCAGCATTTGCT 4603
QY 1047 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 1066
```

```
Db 4604 GAGAACATTCCTATGGAGACAAACAGCCGGTGGTGTTCACAGGAAGAGATCGTGGGGCA 4663
QY 1067 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 1086
Db 4664 GCAAGAGGAGCCAAACATACATGCCCTTCATCGAGTCACTGCCCTAATAATATAGCACTAA 4723
QY 1087 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 1106
Db 4724 GTAGGAGACAAAGGAATCAGCTCTCTGTTGGCCAGAAACAACGCATTCCTAGCTCGT 4783
QY 1107 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 1126
Db 4784 GGCCTTGTGTAGACAGCCTCATATTTTGTCTTTGGATGAAGCCACGCTCCTGGATACA 4843
QY 1127 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1146
Db 4844 GAAAGTGAAGAAGTGTGTCCAAAGCCCTGGCAAGCCCAAGCCAGAGCCGACCTGCAT 4903
QY 1147 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1166
Db 4904 GTGATTTGCTCACCGCTGTCCACCATCCAGATGCAGACTTAATAGTGGTGTTCAGAA 4963
QY 1167 GlyLysIleGlyGluGlnGlyThrHisGlnGlnLeuLeuArgAsnArgAspIleTyrPhe 1186
Db 4964 GGCAGAGTCAAGGAGCATGGCAGCATCAGCAGCTGCTGGCACAGAAAGGAGGCACTAT 5023
QY 1187 LysLeuValAsnAlaGlnSer 1193
Db 5024 TCAATGGTCAGTGTCCAGGCT 5044

RESULT 9
US-09-306-417-2
; Sequence 2, Application US/09306417
; Patent No. US20020103144A1
; GENERAL INFORMATION:
; APPLICANT: Heinrich-Pette-Institut
; TITLE OF INVENTION: Retroviral Gene Transfer Vectors
; FILE REFERENCE: P50491
; CURRENT APPLICATION NUMBER: US/09/306,417
; CURRENT FILING DATE: 1999-05-06
; EARLIER APPLICATION NUMBER: DE 198 22 115
; EARLIER FILING DATE: 1998-05-08
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 8630
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: proviral
; OTHER INFORMATION: plasmid DNA
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(8630)
; OTHER INFORMATION: retroviral expression vector SFbeta91msAI
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(160)
; OTHER INFORMATION: plasmid backbone (pUC)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (161)..(677)
; OTHER INFORMATION: 5'-LTR
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (532)..(1219)
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (1220)..(5062)
; OTHER INFORMATION: mSAI mdrl cDNA
; FEATURE:
; NAME/KEY: misc_feature
```


; LOCATION: (5215)..(5774)
; OTHER INFORMATION: 3'-LTR
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (5775)..(8630)
; OTHER INFORMATION: plasmid backbone (pUC)
US-09-306-417-2

Alignment Scores:

Pred. No.: 1,67e-309 Length: 8630
Score: 3545.00 Matches: 681
Percent Similarity: 75.55% Conservatives: 246
Best Local Similarity: 55.50% Mismatches: 284
Query Match: 58.97% Indels: 36
DB: 10 Gaps: 9

US-09-873-409-6 (1-1195) x US-09-306-417-2 (1-8630)

QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
DB 1370 ATGGTGGTGGAACTTTGGCTGCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 1429
QY 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
DB 1430 GTTTTGGGAAATGACAGATATCTTTGCAATGCGAGGAAATTTAGAGATCTGATGTC 1489
QY 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
DB 1490 AACATCACTAATAGAGTATGATATCAATGATACAGGGTTCCTCATGAATCTGGAGGAAC 1549
QY 44 -----ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
DB 1550 ATGACCAAGGTATGCTTATTATACAGTGAATTTGGTGGGTGCTGGTGTCTGCTTAC 1609
QY 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleGlyGln 81
DB 1610 ATTCAGGTTTCATTTGGTGGCTCGGAGCTGGAGACAAATACACAAATTAAGAAACAG 1669
QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
DB 1670 TTTTTCATGCTATATACGACAGGAGATAGGCTGTTGATGTGCACGATGTGGGGAG 1729
QY 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
DB 1730 CTTAACACCGCACTTACAGATGATGCTCTTAAGATTAATGAAGTATTGGTGACAAAATT 1789
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
DB 1790 GGAATGTTCTTTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGGATTTACAC 1849
QY 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
DB 1850 GGTGGAGACTAACCTTGTGATTTTGGCCATCAGTCTCTGTTCTTGCAGTCTGAGTGTCT 1909
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
DB 1910 GTCTGGGCAAGATACTATCTTCTATTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 1969
QY 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
DB 1970 GGAGCAGTAGCTGAAGAGGCTTTGGCAGCAATTTAGAACTGTGTGATTCATTTGGAGACA 2029
QY 201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
DB 2030 AAGAAGAACTTGAAGGTACACAAATAATTTAGAAAGCTAAAGAATTGGGATTAAG 2089
QY 221 ArgThrIleAlaSerLysValSerLeuGlyValAlaValTyrPhePheMetAsnGlyThrTyr 240
DB 2090 AAAGCTATTACAGCAATATTTCTATAGGTGCTGCTTCTGCTGATCTATGATCTTAT 2149
QY 241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyThrIle 260
DB 2150 GCTCTGGCCTTCTGGTATGGGACCACTTGGTCTCTCTCAGGGAA-----TATTCTATT 2203

QY 261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
DB 2204 GGACAAGTACTCACTGATTTCTTTCTGTATTAAATGGGGCTTTTAGTGTGGCAGGCA 2263
QY 281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
DB 2264 TCTCCAAGCATTTGAAGCATTTGCAATGCAAGGAGGAGCTTATGAAATCTTCAAGATA 2323
QY 301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
DB 2324 ATGTATAATAAGCAAGTATTACAGCTATTGAGAGTGGGACACAAACAGATAATATT 2383
QY 321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
DB 2384 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 2443
QY 341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyLeuThrValAlaLeuValGlyLeu 360
DB 2444 ATCTTGAAGGGCTGAACCTGAAGGTGACAGTGGGACACGGTGGCCCTGGTTGGAAAC 2503
QY 361 AsnGlySerGlyLysSerThrValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
DB 2504 AGTGGCTGTGGAGAGGACACACAGCTCCAGCTGATGCAGAGGCTCTATGCCCCACAGAG 2563
QY 381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
DB 2564 GGGATGGTCAGTGTGTGACGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGAA 2623
QY 401 HisIleGlyValIleValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAnIle 420
DB 2624 ATCATTTGGTGTGGTGTGAGTCCAGAACTGTATTGTTGCCACCACGATAGCTGAAACAT 2683
QY 421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
DB 2684 CGCTATGGCCGTGAAATGTCTCACCATGGATGAGATTGAGAAGCTCTCAAGGAGCCAAAT 2743
QY 441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGly 460
DB 2744 GCCTATGACTTTTATCATGAACCTGCCTCATAAATTTGACACCTGCTGGAGAGAGAGG 2803
QY 461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
DB 2804 GCCCAGTTGAGTGGTGGGAGAGGAGGAGGATCGCATTTGCACGTCGCCCTGGTTGCCAAC 2863
QY 481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
DB 2864 CCCAAGTCTCTGCTGTGATGAGGCCACGTCAGCTTGGACACAGAAAGCGAAGCAGTG 2923
QY 501 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 520
DB 2924 GTTCAGTGGCTCTGGATAAGGCCAGAAAAGGTGCGACCCACCATTTGTGATAGTCTCAT 2983
QY 521 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 540
DB 2984 TTGCTTACACCTTCGTAATGCTGACGCTCATCGCTGTTTCGATGATGGAGTCACTTGGAG 3043
QY 541 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 560
DB 3044 AAGGAAATCATGATGAATCATGAAGAGAAAGGATTTACTTCAAACTGTGCACAAATG 3103
QY 561 Gln-----AspIleLysLysAlaAspGluGlnMetGluSerMetThr 574
DB 3104 CACACAGCAGGAAATGAACCTTGAATTAGAAAATGACGCTGATGAATCCAAAGTGAAT 3163
QY 575 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal----- 589
DB 3164 GATGCTTGAATGCTTCAAAATGATTCAAGATCCAGTCCAGTCTAATAAGAAAAAGATCAACT 3223
QY 590 ---LysSerIleLys---SerAspPheIleAspLysAlaGluSerThrGlnSerLys 607
DB 3224 CGTAGGAGTGTCCGTGGATCACAAAGCCCAAGACAGAAAGCTTAGTACCAAGAGGCTCTG 3283

ORGANISM: RATTUS RATTUS
US-09-769-097-3

Alignment Scores:

Pred. No.: 1,6e-309 Length: 4425
Score: 3540.50 Matches: 680
Percent Similarity: 75.53% Conservative: 243
Best Local Similarity: 55.65% Mismatches: 266
Query Match: 58.89% Indels: 33
DB: 10 Gaps: 9

US-09-873-409-6 (1-1195) x US-09-769-097-3 (1-4425)

```
QY 1 MetIleuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
DB 499 ATGCTGCTGGAACTCTGGCGGCATTATCCATGGAATTCGCTCCCACTTATGATGCTG 558
QY 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
DB 559 GTCTTTGGAGACATGACAGATAGCTTTTGGCAAT-----GTAGGAACAACCGTAGTAGT 612
QY 41 SerPhePheArgLeuThr----- 46
DB 613 AGTTTCTACATCTACAGACATATATGCCAAGCTGGAGGACGAATATGACCGTAGGCC 672
QY 47 LeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeu 66
DB 673 TACTATTACAGGGCATTTGGTGGCGGTGCTCATCGTTGCCCTACATCCAGGTTTCCACT 732
QY 67 TrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGlnPhePheHisSerVal 86
DB 733 TGGTGCTGCGACCTGGGACAAAATACACAAGATTAGGCAGAAAGTTTTTCCATGCCATC 792
QY 87 LeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGluLeuAsnThrArgMet 106
DB 793 ATGAATCAGAGATAGCTGTTTGGTCGTCATGATGAGTGGGAGGCTCAACACCCGGCTC 852
QY 107 Thr---AspIleAspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGln 125
DB 853 ACAGATGACGCTCCAAAATTAATGAAGGAATGGTGACAAAATTTGGAATGTTCTTTTCAG 912
QY 126 AsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLysGlyTyrLysLeuThr 145
DB 913 GCAATGGCAACATTTTTTGGTGGTGTATTAATAGGAATTTACTCGCGGCTGGAAAGCTAAC 972
QY 146 LeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAlaCysSerArgMet 165
DB 973 CTGTGATTTGGCCATCAGCCCTGTTCTTGGACTGTGACTGTGATTTGGGCAAGATA 1032
QY 166 ValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGlu 185
DB 1033 TTGCTCTTCAATTTACTGATAAGGAACCTCAGGCTTATGCAAAAGCTGGAGCAGTTGCTGAA 1092
QY 186 GluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGln 205
DB 1093 GAAGTCTTAGCACCATCAGAACTGTGATTCCTTTGGAGGACAAAAGGAAGAACTTGAA 1152
QY 206 ArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSer 225
DB 1153 AGGTACAATAACAAATTTGGAAAGAGCTTAAAGGCTTGGGATAAAGAAAGCTATCACGGCC 1212
QY 226 LysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTyr 245
DB 1213 AACATTTCCATGGGTGAGCTTTTCTGCTTATCTATGTCATCATATGCTCTGGCATCTCG 1272
QY 246 TyrGlyThrSerLeuLeuLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAla 265
DB 1273 TATGGGACTTCCTTGGTTCATCTCAAAAGAA-----TACACTATTGGACAAGTGTCTCACT 1326
QY 266 ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGlu 285
DB 1327 GTCTTTTCTGTATTAATTTGGAGCATTCAGTGTGGGGCAGGCATCTCCAAATATTGAA 1386
```

```
QY 286 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysIysPro 305
DB 1387 GCCTTTGCCCAATGCTTAGAGGACGCTTATGAAGT---:---:---:---:---:---:---: 1446
QY 306 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 325
DB 1447 AGTATAGACAGCTTCTCAAGAGTGGGCACAAACCCGACCAACATACAGAAATTTGGAA 1506
QY 326 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu 345
DB 1507 TTCMAAAATATTCACTTCACTTACCGTCTCGAAAAGACGTTTCAGATCTTGAAGGCGCTC 1566
QY 346 AsnLeuArgIleLysSerGlyGlnThrValAlaLeuValGlyLeuAsnGlySerGlyLys 365
DB 1567 AACCTGGAAGTGAAGAGCGGCGAGACGTTAGCCCTGGTTGGCAACAGTGGCTGTGGAAA 1626
QY 366 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 385
DB 1627 AGCACAACTGTTCAGCTGCTGAGAGGCTCTACAGCCCATAGAGGCGGAGGTCAGTATC 1686
QY 386 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal 405
DB 1687 GACGGACAGACATCAGGACCATCAATGTGAGGTATCTCGGGGAAATCATTTGGGTGGTG 1746
QY 406 SerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAsp 425
DB 1747 AGTCAGAAACCGCTGCTGTTGCCACCAATTTGCCGAAAACATTCGCTATGCCCGAGAA 1806
QY 426 AspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIle 445
DB 1807 AACGTCACTGATGATGAGAAAGCTGTCAAGAAAGCCCAATGCCATGATTTTCATC 1866
QY 446 MetGluPheProAsnLysPheAsnThrLeuValGlyLysGlyAlaGlnMetSerGly 465
DB 1867 ATGAACCTCCCCACAAATTTGCACCCCTGGTTGGTGAGAGGGGCGCAGCTGAGTGG 1926
QY 466 GlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIle 485
DB 1927 GGACAGAAACAGAGGATCCCATTTGCCGGGCGCTGTGTCGCAACCCCAAGATCTCTTTT 1986
QY 486 LeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaLeu 505
DB 1987 TTGATGAGCGCCACGCTCAGCTTGGACAGAAAGGAAAGCGCTGGTTTCAGGCGGCTCTG 2046
QY 506 GluLysAlaSerLysGlyArgThrThrIleValAlaHisArgLeuSerThrIleArg 525
DB 2047 GATAAGCTTAGAAGGCGCGACCATTTGTGATAGCTCACCGCTTGTCTACAGTTCCG 2106
QY 526 SerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAla 545
DB 2107 AATGCTGATGCTGCTGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2166
QY 546 GluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln-----MetThrTyr 561
DB 2167 GAGCTCATGAGAGAAAGAAATTTACTTCAAACTTGTCACTGACCTCAGACAGCAGGAAAT 2226
QY 562 AspIleLysAlaAspGluGlnMetGluSer-----MetThrTyr 575
DB 2227 GAAATTTGAATTAGAAATTAAGCTTGTGAATCTAAAGATGGAATGATATGTTGGACATG 2286
QY 576 SerThrGluArgLysThrAsnSerLeu-----ProLeuHisSerValLysSerIleLys 593
DB 2287 TCTTCAAAAGATTTCGGATCCAGTCTTAATAAGAAAGATCAACTCGCAAAAGCATCCGT 2346
QY 594 SerAspPheIleAspLysAlaGluSerThr---GlnSerLysGluIleSerLeuPro 612
DB 2347 GGGCCACATGATCAAGACGGGAACTTAGCACCAAGAGGCTCTGGATGACGACGTAACCT 2406
QY 613 GluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTyrProPheValValLeu 632
DB 2407 CCAGCTTCTTTTGGCGGATCTTGATTTGAATTTCACTGAATGGCTTATTTTGTGGTT 2466
QY 633 GlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePheAla 652
```



```
Db 2467 GGTGTAATTTGTCGCAATAAATGAGGCTTGCACACGACATCTCCATAATATTTCA 2526
QY 653 LysIleIleThrMetPheGlyAsnAsnAspLysThrThrLeu---LysHisAspAlaGlu 671
Db 2527 AAGGTTGTAGGGGTTTTACAAAATAATGACACCCCTGAAATCCAGCGGAGAACACAC 2586
QY 672 IleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGln 691
Db 2587 TTGTTTCTTATTTGTCCTGATCCTTGGGATCATCTCTTTTCAATAGTTTTCCTTCAG 2646
QY 692 GlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPhe 711
Db 2647 GGCTTCACATTTGGCAAGCTGGAGAGATCCTCACAAGGAGACTCCGATACATGCTCTTC 2706
QY 712 LysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGluLysGluAsnSerThrGlyGly 731
Db 2707 AAATCCATGCTGAGACAGGACATAAGCTGTTGATGACCCCTAAACACACACAGAGCG 2766
QY 732 LeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIle 751
Db 2767 CTGACACACAGGCTTGCCATGAGCTGCTCAAGTGAAGGGGCTACAGGGTCTAGGCTT 2826
QY 752 GlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyr 771
Db 2827 GCTGTTATTACCCAGAACATAGCAATCTTTGGGACAGGACATCATCTCCCTGATCTAC 2886
QY 772 GlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMet 791
Db 2887 GGCTGGCAATTGACACTTTTACTCTAGCAATTTGTTCCCATCATTTGCTATAGCAGAGT 2946
QY 792 IleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAla 811
Db 2947 GTTGAATGAATGTTGCTGCAACAGCCCTGAAGATGAAGAACTAGAAAGGTTCT 3006
QY 812 GlyLysIleAlaThrGluAlaLeuGluAsnIleAargThrIleValSerLeuThrArgGlu 831
Db 3007 GCGAAGATCGCTACAGAAGCAATTTGAAAACCTTTCGCACCTCTGCTCTTTGACTCGGAG 3066
QY 832 LysAlaPheGluMetTyrGluLysMetLeuGlnThrGlnHisArgAsnThrSerLys 851
Db 3067 CAGAAGTTTGAAATATGATATGCCCAGAGCTGCGAGATCCATACAGAAATGCTTCAAG 3126
QY 852 LysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyr 871
Db 3127 AAGCGCAGCTCTTGGGATCAGCTTCTCTTCCACGAGCCATGATGATTTCTCTAT 3186
QY 872 AlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGly 891
Db 3187 GCTGCTGTTTCCGGTTTGATGCTACTTGTGTGCGACGAGAACTCATGACATTTGAAAT 3246
QY 892 MetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuVal 911
Db 3247 GTTCTGTATTATCTCAGCTATTGCTTTGTTGGTGCACGTGCGAGTGGGCGAGTCAAGTCA 3306
QY 912 LeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuGlu 931
Db 3307 TTGCTCTGACTACGCGAAAGCCAAAGTCTCGCATCTCCCATCATCATGATCATGAG 3366
QY 932 LysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGly 951
Db 3367 AAATCCCTGAGATTGACAGCTACAGCAGCGAGGCTTGAAGCTAATATGTTGGAAGGA 3426
QY 952 AsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeu 971
Db 3427 AATGTGAATTTAATGAGTCACTGTTCAACTATCCACCCGAGCCCAACATCCAGTCTT 3486
QY 972 ArgGlyLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGly 991
Db 3487 CAGGGGCTGAGCCTAGAGGTGAGAAAGGCGACGCTGCGCCCTCGTGGGACAGTGGC 3546
QY 992 CysGlyLysSerThrSerValGlnLeuGlnAlaGluLeuTyrAspProValGlnGlyGln 1011
```

```
Db 3547 TCGGGAGAGGATACAGTGGTCCAGCTGCTTGGCGGCTTCTATGACCCCAATGCGCGAAC 3606
QY 1012 ValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTyrLeuArgSerGlnIle 1031
Db 3607 GGTGTTCTAGATGGCAAGAAATAAAGCAACTCAATGTCCAGTGGCTCCGCGCCACCTG 3666
QY 1032 AlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyr 1051
Db 3667 GGCATTGTGCTCCAGGAGCCCATCTGTTTGTAGCTGCAGCATCGCGAGAACATTCCTAC 3726
QY 1052 GlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAsn 1071
Db 3727 GGAGACAAACAGCCGTGCTGCTCTCATAGAGAGATCGTGAAGGCAGCCCAAGGAGCC 3786
QY 1072 IleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGly 1091
Db 3787 ATCCACAGTTTATGACTCATCTGCTGAGAAATACACACAGATGGGAGCAAGGG 3846
QY 1092 AlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuGlnLys 1111
Db 3847 ACTCAGCTGTCGGGGGCGAGAGCAGCGCATCGCCATCGCGCGCCCTCGTCAGACAG 3906
QY 1112 ProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysVal 1131
Db 3907 CCTCACATCTTACTTCTGGATGAAGCGACATCATCGCTCTCGATACGAGAGTGAAGGTC 3966
QY 1132 ValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArg 1151
Db 3967 GTCCAGAAAGCGCTGCAAAAGCCAGGAGGAGCGCACCTGATGTGTATGTCGCGACCCG 4026
QY 1152 LeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysLysGlu 1171
Db 4027 CTGTCCACCATCCAGAACCGACATTCGTGCTGATTCAGAACGCGCCAGTCAAGGAG 4086
QY 1172 GlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAla 1191
Db 4087 CACGGCACCCACAGCAGCTGCTGGCCCAAGAGCATCTATTTCTCGATGGTCAAGTGTG 4146
QY 1192 GlnSer 1193
Db 4147 CAGGCT 4152

RESULT 11
US-09-769-097-1
; Sequence 1, Application US/09769057
; Patent No. US20020055129A1
; GENERAL INFORMATION:
; APPLICANT: Kimberly Anne Brun
; APPLICANT: Richard James Chenery
; APPLICANT: Harma Ellens
; APPLICANT: John Anthony Feild
; APPLICANT: Lin Yue
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES
; FILE REFERENCE: GP-50009-C2
; CURRENT APPLICATION NUMBER: US/09/769,097
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 09/208,809
; PRIOR FILING DATE: 1998-12-09
; PRIOR APPLICATION NUMBER: 09/156,800
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: US99/20770
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 4369
; TYPE: DNA
; ORGANISM: RATTUS RATTUS
US-09-769-097-1
Alignment Scores: 2.39e-309 Length: 4369
Pred. No.:
```


Score:	3338.50	Matches:	680
Percent Similarity:	75.53%	Conservative:	243
Best Local Similarity:	55.65%	Mismatches:	266
Query Match:	58.86%	Indels:	33
DB:	10	Gaps:	9

US-09-873-409-6 (1-1195) x US-09-769-097-1 (1-4369)

Qy	1	MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu	20
Db	499	ATGCTGTGGGAACCTCGCGGCCATTATCCATGGAATTCGCTCCCACTTATGATGCTG	558
Qy	21	ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr	40
Db	559	GTCTTTGGAGATGACAGATAGCTTTGCANAAT-----GTAGGAACAACCGTAGTATG	612
Qy	41	SerPhePheArgLeuThr-----	46
Db	613	AGTTTCTACAAATGCTACAGACATATATGCCAAGCTGGAGAGCAAAATGCCACGCTACGCC	678
Qy	47	LeuTyrTyrValGlyIleGlyValAlaIleLeuPheGlyTyrIleGlnIleSerLeu	66
Db	673	TACTATTACAGCGCAATGTGCGGTGTGCTCATCGTTGCCCTACATCCAGGTTTCACTT	732
Qy	67	TrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGlnPhePheHisSerVal	86
Db	733	TGGTGCCTCGCAGCTGGGAGCAAAATACACAAGATTAGGCAGAGATTTTTTCCATGCCCAT	792
Qy	87	LeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyLeuLeuAsnThrArgMet	106
Db	793	ATGAATCAGAGATAGTCTGGTTTGACGTGCATGACGTTGGGGAGCTCAACACCGGCTC	852
Qy	107	Thr---AspIleAspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGln	125
Db	853	ACAGATCAGCTCCAAAATTATGAAGAAATGGTGACAAAATTTGGAATGTTCTTTTCAG	912
Qy	126	AsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLysGlyTrpLysLeuThr	145
Db	913	GCAATGGCAACATTTTTTGTGTGTTTTATAATAGGATTTTACTCGCGGGTGGAAAGCTAACT	972
Qy	146	LeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAlaCysSerArgMet	165
Db	973	CTTGATGATTTGGCCATCAGCCCTGTTCTTGGAGCTGTACGCTGGTATTGGGCAAGATA	1032
Qy	166	ValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGlu	185
Db	1033	TTGTCTTCAATTACTGATAGGAACCTCAGGCTTATGCAAAAGCTGGAGCAGTTGCTGAA	1092
Qy	186	GluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGlyLysGlnLeuGln	205
Db	1093	GAAGCTTTAGCAGCCATCAGAACTGTGATTCCTTTGGAGAGCAAAAGAGAACTTGAA	1152
Qy	206	ArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSer	225
Db	1153	AGGTACAAATTAACAAATTTGGAAGAGCTAAAGAGCTTGGGATTAAGAAGCTATACAGGCC	1212
Qy	226	LysValSerLeuGlyAlaValTyrPhePheMetAsnGlyTyrTyrGlyLeuAlaPheTrp	245
Db	1213	AACATTTCCATGGTGGCAGCTTTCTGCTTATCTATGATCATATGCTCTGGCATTCGG	1272
Qy	246	TyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAla	265
Db	1273	TATGGAGACTTCTTGGTCATCTCAAAAGAA-----TACACTATTGGCAAGTGCTCACT	1326
Qy	266	ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGlu	285
Db	1327	GTCTTTTTTCTGPTAATTATGGAGCATTAACGTTGGTCAGCGCATCTCCAAATATTGAA	1386
Qy	286	ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPro	305
Db	1387	GCCTTCGCCAATGCTAGAGAGCAGCTTATGAAGACTTTCAGTATTAATTGATTAATTAAGGCC	1446
Qy	306	SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu	325


```
Db 2527 AAGGTTGTAGGGGTTTTTACAAAAAATGACACCCCTGAATCCAGCGCGCAGAACACAAAC 2586
Qy 672 IleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGln 691
Db 2587 TTTGTTTCTTTATGTTCCGTGATCCCTGGGATCATCTCTTTCAATGAGTTTTTCTCTCAG 2646
Qy 692 GlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisIleAlaPhe 711
Db 2647 GGCTTCACATTGGCAAGCTGGAGAGATCCTCACCAAGGAGCTCCGATACATGTTCTTC 2706
Qy 712 LysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGluLysGluAsnSerThrGlyGly 731
Db 2707 AAATCCATGCTGAGCAGGACATAAGCTGTTGATGACCCCTAAAAACACCCAGGAGCG 2766
Qy 732 LeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIle 751
Db 2767 CTGACCACCAAGGCTTGCCAATGACGCTGCTCAAGTGAAGGGGCTACAGGGTCTAGGCTT 2826
Qy 752 GlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyr 771
Db 2827 GCTGTTATTACCAACATAGCAATCTTGGGACAGGCATCATATCCCTGATCTAC 2886
Qy 772 GlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMet 791
Db 2887 GGCTGGCAATTGACATCTTTACTCTAGCAATTTGTTCCCATCATTTGCTATAGCAGGAGT 2946
Qy 792 IleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAla 811
Db 2947 GTTGAAATGAAATGTTGTCTGGCAAGCGCTGAAGATGAAGAGCACTAGAAAGGTTCT 3006
Qy 812 GlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGlu 831
Db 3007 GGGAGATCGCTACAGAGCAATTTGAAAACCTTTCGCACCTGCTCTTTGACTCGGAG 3066
Qy 832 LysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLys 851
Db 3067 CAGAATTTGAAACTATGTATGCTGCGACAGGCTTGCAGATACCATACAGAAATGCTTTGA 3126
Qy 852 LysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyr 871
Db 3127 AAAGCGACGCTTTTGGGATCACCTTTCTCTTTCACCCAGGCGCATGATGTATTTCTCTAT 3186
Qy 872 AlaAlaGlyPheArgPheGlyAlaTyrIleLeuIleGlnAlaGlyArgMetThrProGluGly 891
Db 3187 GCTGCTGTTTCCGGTTTGATGCTTACTTGTGCGCAGAGAACTCATGACATTTGAAAT 3246
Qy 892 MetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuVal 911
Db 3247 GTTCTGTTAGTATTCTCAGCTATTGTTCTTTGGTCCCATGGCAGTGGGCGCAGGTCA 3306
Qy 912 LeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuGlu 931
Db 3307 TTCGCTCCTGACTACGCAAGCCAAAGTCTCTCGGCATCCCATCATCAGGATCATTTGAG 3366
Qy 932 LysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGly 951
Db 3367 AAAATCCCTGAGATTGACAGTACAGCACGAGGAGGCTTGAAGCCTATATATGTTGGAAG 3426
Qy 952 AsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeu 971
Db 3427 AATGTGAAATTTAATGGAGTTCATGTTCAACTATCCACCCGACCAACATCCAGTGTCT 3486
Qy 972 ArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGly 991
Db 3487 CAGGGGCTGAGCCTAGAGGTGAAGAAAGGAGCAGCGCTGGGCCCTCTGGGGCAGGAGTGG 3546
Qy 992 CysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGln 1011
Db 3547 TCGGGGAAGATACAGTGTCTCAGCTGTTGAGCGCTCTCTATGACCCCATGCGCGGAACA 3606
Qy 1012 ValLeuPheAspGlyValAlaAspAlaLysGlnLeuAsnValGlnTyrLeuArgSerGlnIle 1031
Db 3607 GTGTTTCTAGATGGCAAGAAATAAGCAACTCAATGTCCAGTGGCTCCGCGGCCACCTG 3666
```

```
Qy 1032 AlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyr 1051
Db 3667 GGCAATTGTCTCCAGGAGCCCATCTGTTTGTAGTCTGAGCATCGCCGAGAACATTTGCTTAC 3726
Qy 1052 GlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAsn 1071
Db 3727 GGAGACAAACAGCCGTTGCTGCTCTCATAGAGAGATCGTGAAGGCGACCAAGGAGGCAAC 3786
Qy 1072 IleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGly 1091
Db 3787 ATCCACCACTTTCATGACTCATCTGCTGAGAAATACAAACACAGAGTGGGAGACAAAGGG 3846
Qy 1092 AlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLys 1111
Db 3847 ACTCAGCTGTCGGGGGCGAGAACGAGCATCGCCATCGCGCGCCCTCTGTCGACAGAC 3906
Qy 1112 ProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysVal 1131
Db 3907 CCTCACATCTTACTTCTGGATGAAGCGACATCAGCTCTCGATACCGAGAGTGAAGAGTTC 3966
Qy 1132 ValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArg 1151
Db 3967 GTCAGGAAGCGCTGGACAAAGCCAGGAGGCGCCACCTGCAATTGTGATCGCGCACCGC 4026
Qy 1152 LeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGlu 1171
Db 4027 CTGTCCACCATCCAGAACCGACAGACTTGTGCTGTTGATTTCAGAACCGCCAGGTCAAGG 4086
Qy 1172 GlnGlyThrHisGlnGluLeuArgAsnArgAspIleTyrPheLysLeuValAsnAla 1191
Db 4087 CACGGCACCCACAGCAGCTGCTGGCCCAAGAAAGGCATCTATTTCTCGATGGTTCAGTGT 4146
Qy 1192 GlnSer 1193
Db 4147 CAGGCT 4152
```

RESULT 12

```
US-10-072-621-2
; Sequence 2, Application US/10072621
; Patent No. US20020169137A1
; GENERAL INFORMATION:
; APPLICANT: Reiner, Peter B.
; APPLICANT: Connop, Bruce P.
; APPLICANT: Pollard, Michelle
; TITLE OF INVENTION: REGULATION OF AMYLOID PRECURSOR PROTEIN EXPRESSION
; TITLE OF INVENTION: BY MODIFICATION OF ABC TRANSPORTER EXPRESSION OR ACTIVITY
; FILE REFERENCE: 100103.402
; CURRENT APPLICATION NUMBER: US/10/072,621
; CURRENT FILING DATE: 2002-02-08
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 4643
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-072-621-2
```

```
Alignment Scores:
Pred. No.: 3,21e-309 Length: 4643
Score: 3537.50 Matches: 681
Percent Similarity: 75.55% Conservative: 246
Best Local Similarity: 55.50% Mismatches: 263
Query Match: 58.84% Indels: 37
DB: 9 Gaps: 10
```

US-09-873-409-6 (1-1195) x US-10-072-621-2 (1-4643)

```
Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 575 ATGGTGTGGGAACCTTTGGCTGCCATCATCCATGGGGCTGGACTTCTCTCATGATGCTG 634
Qy 21 ValLeuGlyLeuMetSerAsp-----AsnLeuIleSer 31
```


[illegible]

1706	AGTGGCTGTGGGAAGACCAACAGTCTCAGCTGTATGTGCAGAGGCTCTATGACCCACACAGAG	1765
381	GlyPheIleMetValAspGluAenAspIleAArgAlaLeuAenValAArgHisTyArgAsp	400
1766	GGGATGGTCACTGTTGATGGCAGGATATTAGGACCATATAATGTAAAGTTTCTACGGGA	1825
401	HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle	420
1826	ATCATTTGTTGGTGTAGTCAGGAACCTGTATTGTTTGGCCACCACCATGAGCTGAAACATT	1885
421	LysTyArgArgAspValThrAspGluGluMetGluAArgAlaAalaArgGluAalaen	440
1886	CGCTATGGCCGTGAAATATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGAGAGCCCAAT	1945
441	AlaTyArgPheIleMetGluPheProAsnLysPheAsnThrLeuValClyGluLysGly	460
1946	GCCTATGACTTTATCATGAACCTGCCTCATTAATTTGACACCTGGTTGGAGAGAGAGG	2005
461	AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValAArgAsn	480
2006	GCCCAAGTTGATGTGGCGCAGAACAGAGGATCGCATTTGACAGTGTGCCCTGGTTCGCAAC	2065
481	ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla	500
2066	CCCAAGATCTCTCTGTGGATGAGGCCAGTCAGCTTGGACAGAAGCGAAGAGCTGT	2125
501	ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg	520
2126	GTTTCAGGTGGCTCTGGATAGGCCAGAAAGGTCGGACCCACCATTTGTATAGCTCATCGT	2185
521	LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu	540
2186	TTGTCTCAGTTCTGTAATGTCGATCGTCATCGCTGGTTTCGATGATGAGAGTCATTGTGGAG	2245
541	LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyTySerLeuValMetSer	560
2246	AAAGAAATCATGATGAACTCATGAAGAAGAGGCAATTTACTTCAAACTTTGTCACAATG	2305
561	Gln-----AspIleLysLysAlaAspGluGlnMetGluSerMetThr	574
2306	CAGACAGCGAAGTAAGTTGAATTAGAAATGCAGCTGATGAATCCAAAGTGAATTT	2365
575	TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal-----	589
2366	GATGCTTGGAAATGCTCTCAAAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACT	2425
590	---LysSerIleLys---SerAppPheIleAppLysAlaGluLysSerThrClnSerLys	607
2426	CGTAGGAGTGTCCCGTGATCACAAGCCCAGACCTTAGTAGTACCAAGAGGCTCTG	2485
608	GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrp	627
2486	GATGAAGTATACCTCAGTTTCTTTTGGAGGATATGAGACTAAATTTAACTGNAATGG	2545
628	ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe	647
2546	CGTTATTTTGTGTTGTTGTTATTTTGTGCCATTATATAATGGAGGCTGTCAACACGACATT	2605
648	SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu	666
2606	GCAATATATTTTCCAAAGATTAATAGGGGTTTTTACAAAGAAATTTGATGAATCTGGAACAAA	2665
667	LysHisAspAlaGluIleTySerMetIlePheValIleLeuGlyValIleCysPheVal	686
2666	CGACAGATAGTACTGTTTTCACCTATGTTTCTAGCCCTTGGAAATTAATTTCTTTTATT	2725
687	SerTyPheMetGlnGlyLeuPheTyGlyArgAlaGlyGluIleLeuThrMetArgLeu	706
2726	ACATTTTTTCTCAGGCTTTTCAATTTGGCAAGCTGGAGAGATCCTCACCAAGCGGCTC	2785
707	ArgHisLeuAlaPheLysAlaMetLeuTyArgIleAlaTrpPheAspGluLysGlu	726
2786	CGAATCATGGTTTTTCCCATGCTCAGACAGATGTGAGTTGGTTTGTATGACCTTAAA	2845


```
QY 727 AspSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyVala 746
|||
Db 2846 AACCACTCGAGCATGCTACAGCTCGCAATGATGCTCTCAGTTAAAGGGCT 2905
|||
QY 747 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValile 766
|||
Db 2906 ATAGTTCCAGGCTGCTGTAATATCCAGATATAGCAATCTTGGCAGCAATAATT 2965
|||
QY 767 IleSerPheIleTyThrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 786
|||
Db 2966 ATATCTTCTATGCTGTCGCAACTAACACTGTTACTCTTAGCAATTGTACCATCAT 3025
|||
QY 787 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 806
|||
Db 3026 GCAATAGCAGGAGTGTGTGAATAATGTTGTCTGAGCAAGCAGCTGGAAGATAAGAAA 3085
|||
QY 807 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 826
|||
Db 3086 GAACAGAGGCTGCGGAAGATCGCTACTGAGCAATAGAAAACTTCCGAAACCGTTGT 3145
|||
QY 827 SerLeuThrArgGluLysAlaPheGluGlnMetTyThrGluMetLeuGlnThrGlnHis 846
|||
Db 3146 TCTTTGACTCAGGAGCAGAAAGTTTGAACATATGATGCTCAGAGATTTCAGAGTACCATAC 3205
|||
QY 847 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyThrAlaPheSerHisAlaPhe 866
|||
Db 3206 AGAACTCTTTGAGAAAGCAGACATCTTGGAAATTACATTTTCCTTACCCAGGCATG 3265
|||
QY 867 IleTyThrPheAlaTyAlaAlaGlyPheArgPheGlyAlaTyThrLeuIleGlnAlaGlyArg 886
|||
Db 3266 ATGTATTTTCTATGCTGGATGTTTCGGTTTGGAGCTTACTTGGTGCCACATAAACTC 3325
|||
QY 887 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyThrGlyAlaMetAlaIle 906
|||
Db 3326 ATGAGCTTTCAGGATGTTCTGTAGTATTTTTCAGCTGTTCTTGTGTGCTGCGCGTG 3385
|||
QY 907 GlyLysThrLeuValLeuAlaProGluTyThrSerLysAlaLysSerGlyAlaAlaHisLeu 926
|||
Db 3386 GGGCAAGTCAGTTCATTTGCTTCCTGACTATGCCAAAGCCAAATATCAGAGGCCACATC 3445
|||
QY 927 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 946
|||
Db 3446 ATCATGATCATGAAACCCCTTTGATTGACAGCTACAGCAGGAAGCCCTAATCCCG 3505
|||
QY 947 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPheTyThrProCysArgPro 966
|||
Db 3506 AACACATTGGAAGGAAATGTCCATTTTGGTGAAGTTGTATTCAACTATCCACCCGACCG 3565
|||
QY 967 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 986
|||
Db 3566 GACATCCAGTGTCTTCAAGGACTGAGCTGAGGTGAAGAGGCCAGACGCTGGCTCTG 3625
|||
QY 987 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyThrAsp 1006
|||
Db 3626 GTGGCAGCAGTGTGTGGGAAGACACAGTGTGTCCAGCTCTCTGGAGCGTCTTACGAC 3685
|||
QY 1007 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 1026
|||
Db 3686 CCCTTGGCAGGAAAGTGTCTTGTATGTGGCAAGAAATAAAGCAGCAATGTTCACTGG 3745
|||
QY 1027 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 1046
|||
Db 3746 CTCGAGCACACCTGGGCATCGTGTCCAGGAGGCCATCTCTGTTGACTGCAGCATTTGCT 3805
|||
QY 1047 GluAsnIleAlaTyThrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 1066
|||
Db 3806 GAGAACATTGCTTATGGAGACAAACAGCCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3865
|||
QY 1067 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyThrAsnThrGln 1086
|||
Db 3866 GCAAGGAGGGCCACATACATGCTTTCAGAGTCACTGCGCTTAATAATAATAGCACTAAA 3925
|||
```

```
QY 1087 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 1106
|||
Db 3926 GTAGGAGACAAAGGAACCTCAGCTCTCTGGTGCCAGAAACGCAATGCCATAGCTCGT 3985
|||
QY 1107 AlaLeuLeuGlnLysProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspAsn 1126
|||
Db 3986 GCCCTTGTGTAGACAGCTCATATTTGCTTTTGGATGAAGCCACGCTCAGCTCTGGATACA 4045
|||
QY 1127 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1146
|||
Db 4046 GAAAGTGAANAAGTTGTCCAGAGCCCTGGCAAAAGCCAGAGAGCCGCGCCTGCATT 4105
|||
QY 1147 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1166
|||
Db 4106 GTGATTGCTCACCGCTGTCCACCATCCAGATGCAGACTTAATAGTGTGTTTCAGAAAT 4165
|||
QY 1167 GlyLysIleGlyGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyThrPhe 1186
|||
Db 4166 GGCAGAGTCAAGAGCATGGCAGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTT 4225
|||
QY 1187 LysLeuValAsnAlaGlnSer 1193
|||
Db 4226 TCAATGCTCAGTGTCCAGCT 4246
|||
RESULT 13
US-09-866-866A-5
; Sequence 5, Application US/09866866A
; Patent No. US20020102244A1
; GENERAL INFORMATION:
; APPLICANT: Sorrentino, Brian
; APPLICANT: Schuetz, John
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
; FILE REFERENCE: 1340-1-02CIP2
; CURRENT APPLICATION NUMBER: US/09/866, 866A
; CURRENT FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 09/584, 586
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: PCT/US99/11825
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: 60/086, 988
; PRIOR FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 4189
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-866-866A-5
Alignment Scores:
Pred. No.: 2,71e-308 Length: 4189
Score: 3526.50 Matches: 678
Percent Similarity: 74.74% Conservative: 242
Best Local Similarity: 55.08% Mismatches: 262
Query Match: 58.66% Indels: 49
DB: 10 Gaps: 10
US-09-873-409-6 (1-1195) x US-09-866-866A-5 (1-4189)
QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
|||
Db 148 ATGATTCTGGGAACCTCTCGCTGCTATTATCCATGGAACATTACTTCCCTCTTGTATGCTG 207
|||
QY 21 ValLeuGlyGluMetSerAspAsnLeuIle----- 30
|||
Db 208 GTGTTTGGAAACATGACAGATAGTTTACAAAAGCAGAACCCAGTATTCTGCCAAGCATT 267
|||
QY 31 -----SerGlyCysLeuValGlnThrAsnThr----- 39
|||
Db 268 ACTAATCAAAAGTGGACCCCAACAGTACTCTCATCATCAGCAACAGCAGCTCTGGAGGAAGAG 327
|||
QY 40 -----TyrSerPhePheArgLeuThrLeuTyThrValGlyIleGlyValAlaAla 56
|||
```


Db 328 ATGCCCATATACCCCTAC-----TATTACACCGGGATGGTGTGTGTG 372
Qy 57 LeuIlePheGlyTyrIleGlnIleSerLeuTrrPheIleThrAlaIaArgGlnThrLys 76
Db 373 CTCATAGTTGCCCTACATCCAGGTTTCACITTTGGTGGCTGGCAGCTGGAACACAGATACAC 432
Qy 77 ArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrrPheAspSer 96
Db 433 AAGATTAGGCAGAAAGTTTTCATGCTGATAATGAATCAGGAGATAGGCTGGTTTGATGTG 492
Qy 97 CysAspIleGlyGluLeuLeuThrArgMetThr--AspIleAspLysIleSerAspGly 115
Db 493 CATGATTGTTGGGAGCTCAACACCGGCTCACAGATGATGTCTCCAAATTAATGACGGA 552
Qy 116 IleGlyAspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAla 135
Db 553 ATGTGTGACAAATTTGGGATGTTTTTCAGTCCATAACACACATTTTAGCCGTTTATC 612
Qy 136 ValGlyLeuValLysGlyTrrPysLeuThrLeuValThrLeuSerThrSerProLeuIle 155
Db 613 ATAGGATTTATAAGTGGTTGGAAGCTAAACCTTGTCAATTTGGCTGCAGCCCTCTTATT 672
Qy 156 MetAlaSerAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSer 175
Db 673 GGATTCATCTCTTTGTGGGCAAGGATTGACTTCAATTTACTAATAAGGAACCTCCAG 732
Qy 176 AlaTyrSerLysAlaGlyAlaValAlaGluValLeuSerSerIleArgThrValIle 195
Db 733 GCTTATGCAAAAGCTGGAGGATGCTGGAAGAAGTCTTAGCAGCCATCAGAACTGTGATT 792
Qy 196 AlaPheArgAlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLys 215
Db 793 GCCTTTGGAGGCAACACAGAGGAACCTGAAAGGTGCAATAAATAATTAGAAAGCTAAA 852
Qy 216 AppPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePhe 235
Db 853 AATGTTGGCATAAAGAACGTATCACAGCCAGCATTTCCATAGGCATTCCTACCTGTTG 912
Qy 236 MetAsnGlyThrTyrGlyLeuAlaPheTrrPyrGlyThrSerLeuIleLeuAsnGlyGlu 255
Db 913 GTCTATGCATCATATGCATGGCATTCGTGTATGGACATCTTGGTCTCTCAATGAA 972
Qy 256 ProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerTyr 275
Db 973 -----TATTCTATTGGAGAAGTGTACTGTCTCTCTCTATTGTTGGGCACTTT 1026
Qy 276 CysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPhe 295
Db 1027 AGTATTGGACACTTGGCCCAACATAGAACCTTTGCAAAACGACGAGGGGCGACTTT 1086
Qy 296 HisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyr 315
Db 1087 GAATCTTCAAGATAATTGATAACGAGCCNAGCATTTGACAGCTTCTCAACAAAGGCGTAC 1146
Qy 316 LysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSer 335
Db 1147 AAACACAGACAGTATAATGGGAACCTTAGAGTTTAAAAATGTTCCACTTCAACTACCCATCG 1206
Qy 336 ArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrVal 355
Db 1207 AGAAGCGAAATTCAGATCTTGAAGGGCCCTCAATCTGAAGGTGAAGCGGACAGCGGTG 1266
Qy 356 AlaLeuValGluLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeu 375
Db 1267 GCCTTGGTGGCAACAGTGGCTGTGAAAAAAGCACAACTGTCTCCAGCTGTAGTGCAGAGGCTC 1326
Qy 376 TyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnVal 395
Db 1327 TAGACCCCTCGAGGGCGTGGTCAGTATCGACGGACAGACATCAGACCATCAATGTG 1386
Qy 396 ArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThr 415
Db 1387 AGGTATCTGAGGGAGATCATTTGGTGTGGTGTGAGTCAAGAACTGTGTGTTTGGCCACG 1446

Qy 416 IleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAla 435
Db 1447 ATGCCCGAGAAACATTTCGCTATGCGCGAGAAGATGTCCACCATGATGAGATTGAGAAAGCT 1506
Qy 436 AlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeu 455
Db 1507 GTCAAGGAAGCAACATGCCCTATGACTTCATCAATCAATGCAATTTGACACCCCTG 1566
Qy 456 ValGlyGluLysGlyValaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArg 475
Db 1567 GTTGTGAGAGAGGGCGCGAGCTGAGTGGGGACAGAAACAGAGAATCCCATTTGCCCGG 1626
Qy 476 AlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSer 495
Db 1627 GCCCTGGTCCGCAATCCCAAGATCCTTTGTTGGAGCGCCACCTCAGCCCTGGATACA 1686
Qy 496 GluSerLysSerAlaValGlnAlaLeuGluLysAlaSerLysGlyArgThrThrIle 515
Db 1687 GAAAGTGAAGCTGTGTCGACGCCGCACTGGATAAGGCTAGAGAAGCGCGACCACTT 1746
Qy 516 ValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAsp 535
Db 1747 GTGATAGCTCATCGCTTGTCTACAGTTCTGTAATGTGACGTCACTGCTGGTTTGATGTG 1806
Qy 536 GlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyr 555
Db 1807 GGTGTGATTCGACCAAGGAATCATGATGAGCTCATGAGAGAAAGGCGCATTTACTTC 1866
Qy 556 SerLeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyr 575
Db 1867 AAACTTGTCTATGACACAG---ACTAGAGGAATGAAATTTGAACCCAGGAATAATGCTTAT 1923
Qy 576 SerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAsp 595
Db 1924 GGATCCAGAGTGACACTGATGCTTCTGAACCTGACTTCAGAAGAAATCC---AAATCACCT 1980
Qy 596 PheIleAspLysAla-----GluGluSerThrGlnSerLysGluIle 609
Db 1981 TTAATAAGGAGATCAATTTACAGAGTGTCCACAGAAACCAAGACCACAGAGAGAACATT 2040
Qy 610 SerLeu-----ProGluValSerLeuLeuLysIleLeuLys 621
Db 2041 AGTATGAAAGAGCGCTGTGGATGAAGATGTGCTCTGTTTTCCTTTTGGCGGATCTTAAAT 2100
Qy 622 LeuAsnLysProGluTrrProPheValValLeuGlyThrLeuAlaSerValLeuAsnGly 641
Db 2101 CTAAATCTAAGTGAATGGCTTATTACTTGTGGCTACTTTGCGCTGTTTATAAATGGG 2160
Qy 642 ThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGly---Asn 660
Db 2161 TGCATACACACAGTGTTCCTCATAGTATTTTCAAGGATGTAGGGGTTTTTCAAGAGAT 2220
Qy 661 AsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeu 680
Db 2221 GATGACCATGAAACTAAACACAGAAATTGTAATTTGTTTCTCTGTTCTTCTGTTATG 2280
Qy 681 GlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGlu 700
Db 2281 GGGCTGATTTCTTTTGTATATATTCTTTTTCAGGGGCTTCACATTTGGCAAGCCGAGAG 2340
Qy 701 IleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAla 720
Db 2341 ATCTCACCAACGAGTCCGATACATGTTTTCNAATCCATGCTGACAGAGATATAGC 2400
Qy 721 TrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIle 740
Db 2401 TGGTTCGATGACATAAAGAAACAGCACTGGCTCACTGACCACAGGCTCGCCAGTGATGT 2460
Qy 741 AlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsn 760
Db 2461 TCTAGTGTAAAAGGGGCGATGGCGCGAGGCTGTGTTAGTTATCCCAAGAAATGTAGCAAC 2520


```
QY 761 MetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeu 780
Db 2521 CTCGGACAGGAGTCATCCTCTCTGATGCTATGCTGGCGACCATCTTCTACTTGTGA 2580
QY 781 SerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPhe 800
Db 2581 GTAATTATACCGCTCAVTTGTTGGCGGAATTATTGAATGAAGCTGTTGTCTGCCAA 2640
QY 801 AlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGlu 820
Db 2641 GCCTTGAGGACAAAGAACAGCTTGAGATCTCTGGGAGAGATTGCTACAGAAGCAATTGAA 2700
QY 821 AsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGlu 840
Db 2701 AACTTCGCACTATTGTCTCTTTGACTCGGAGCAAGATTTGAAACCAATGATGCGCCAG 2760
QY 841 MetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyr 860
Db 2761 AGCTTGAGGTACCATACAGAAATGCGATGAAGAAAGCACACGTGTTGGGATCAGCTTC 2820
QY 861 AlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyr 880
Db 2821 TCCTTACCAGCCCATGATGATTTTCTATGCTGCTTTCGGTTCGGTGCTCTAC 2880
QY 881 LeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAla 900
Db 2881 TTGCTGSCACAAACACTCATGACTTTTGAAGATGTTATGTTGTTATTTCTGCTGTGTC 2940
QY 901 TyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLys 920
Db 2941 TTTGGTGCCATGCGAGCTGGGAACTACTAGTTTCATTTGCTCTGACTATGCGAAGCCAAA 3000
QY 921 SerGlyAlaAlaHisLeuPheAlaLeuGluLysLysProAsnIleAspSerArgSer 940
Db 3001 GTATCAGCTCTCATATCATCAGGATCTTTGAGAAACCCCTGAGATTGACAGCTACAGC 3060
QY 941 GlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhe 960
Db 3061 ACAGAGGCTTGAAGCTCTCTGTTAGAAGAAATGTAATAATTAATGAGTCCAGTTT 3120
QY 961 PheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArg 980
Db 3121 AACTATATCCACCCAGCCACATCCCATGCTTCCAGGGGTGAGCCTCGAGGTGAAGAAG 3180
QY 981 GlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeu 1000
Db 3181 GCCCAGACGTGGCCCTGGTGGGAGCAGTGGCTGTGGGAGAGCACAGTGTCTCCAGCTG 3240
QY 1001 LeuGlnArgLeuTyrAspProValGlnGlnValLeuPheAspGlyValAspAlaLys 1020
Db 3241 CTCGAGCGCTTCTACGACCCCATGGCTGGATCAGTGTCTTAGATGCCAAAGAAATAAG 3300
QY 1021 GluLeuAsnValGlnThrLeuArgSerGlnIleAlaIleValProGlnGluProValLeu 1040
Db 3301 CAACTGAATGTCCAGTGGCTCCGAGCTCACCTTGGCAATTGTGCCAGAGGCCATCTC 3360
QY 1041 PheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeu 1060
Db 3361 TTTGATGTCAGCATGTGAGAGAAATCGCTTATGAGACAAACCCGGCGGTGTCTCAT 3420
QY 1061 AspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuPro 1080
Db 3421 GAGGAGATTGTGAGGCGAGCCAGGAGGCCAACATCCACCAGTTTCATCGACTCAGTGCCT 3480
QY 1081 GluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlnLysGln 1100
Db 3481 GATAAATACAAACACAGAGTAGGAGACAAAGGCACCTCAGCTGTCTGGGTGGCGAAGCAG 3540
QY 1101 ArgLeuAlaIleAlaArgAlaLeuGlnLysProLysIleLeuLeuLeuAspGluAla 1120
Db 3541 CGATCCCATCGCAGCTGCTCTGTCAGACAGCCTCACATTTTACTTCTGAGAGAGCA 3600
QY 1121 ThrSerAlaLeuAspAsnArgSerGluLysValGlnHisAlaLeuAspLysAlaArg 1140
```

```
Db 3601 ACATCAGCTCTGATACAGAAAGTGAAGAGTGTCCAGAAAGCGCTGGACAAAGCCAG 3660
QY 1141 ThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeu 1160
Db 3661 GAAGGCCGACCTGCTATGCTGCTCATCCGCTGTCCACCATCCAGAACGGGACTTG 3720
QY 1161 IleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArg 1180
Db 3721 ATCGTGGTATTGAGAACGGCAAGTCAAGGAGCAGCGCACCCACGACGAGCTGCTGGCG 3780
QY 1181 AsnArgAspIleTyrPheLysLeuValAsnAla 1191
Db 3781 CAGAAGGCGATCTACTTCTCAATGTGTCCAGGCT 3813
RESULT 14
US-09-866-866A-7
; Sequence 7, Application US/09866866A
; Patent No. US2002010244A1
; GENERAL INFORMATION:
; APPLICANT: Schuetz, John
; APPLICANT: Sorrentino, Brian
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
; FILE REFERENCE: 1340-1-021CIF2
; CURRENT APPLICATION NUMBER: US/09/866,866A
; PRIOR FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 09/584,586
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: PCT/US99/11825
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: 60/086,988
; PRIOR FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 4788
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-866-866A-7
Alignment Scores:
Pred. No.: 7,58e-308 Length: 4788
Score: 3522.50 Matches: 668
Percent Similarity: 74.74% Conservative: 252
Best Local Similarity: 54.26% Mismatches: 264
Query Match: 58.59% Indels: 47
DB: 10 Gaps: 8
US-09-873-409-6 (1-1195) x US-09-866-866A-7 (1-4788)
QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 148 ATGCTGGTGGAACTCTGGCTGCTATATCCATGGAGTGGCGCTCCCACTTATGATGCTG 207
QY 21 ValLeuGlyCyluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsn----- 38
Db 208 ATCTTTGGTGACATGACAGATAGCTTTGCAAGTGTAGGAAACGCTCTCTAAAAACAGTACT 267
QY 39 -----Thr 39
Db 268 AATATGAGTGAGSCCGATAAAGAGCCATGTTTGCACAACTGGAGGAAGAAATGACCAAG 327
QY 40 TyrSerPhePheArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePhe 59
Db 328 TAGCCTAC-----TATTACCGGAGATTGGTGTGCTGTGCTCATAGTT 372
QY 60 GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg 79
Db 373 GCCTACATCCAGTTTCATTTTGGTCCCTGGCAGCTGGAAGACAGATACACAAGATCAGG 432
QY 80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99
Db 433 CAGAAGTTTTTTCATCTATAATGAATCAGGATAGGCTGGTGTGATGTGATGATGATGACGTT 492
```


Qy	100	GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp	118
Db	493	GGGAGAGCTCAACCCCGGCTCAGATGATGTTTCCAAAATTAATGAAGNAATTGGTGAC	552
Qy	119	LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu	138
Db	553	AAATCGGAATGTTCTTCAGGCAATGGCAATTTTTTGGTGGTTTTTAATAGGATTT	612
Qy	139	ValysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer	158
Db	613	ACCGTGGCTGGAAGCTAAACCTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTGCA	672
Qy	159	AlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer	178
Db	673	GCTGGTATTTGGGCAAGAATATTGTTCTCATTTACTGATGAAGCACTCCATGCTTATGCA	732
Qy	179	LysAlaGlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArg	198
Db	733	AAAGCTGGACAGTGTCTGAGNAGTCTTAGCAGCCTCAGACATGATGTTGCTTTGGA	792
Qy	199	AlaGlnLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly	218
Db	793	GGACAAAGAAGAACTTGAAGGTACAATAACAACCTTGGAGAAGCTTAAAGGCTGGG	852
Qy	219	IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGly	238
Db	853	ATAAAGAAAGCTATCAGCGCAACATCTCCATGGGTGAGCTTTTCTCTTATCATGCA	912
Qy	239	ThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr	258
Db	913	TCATATGCTCTGGCATCTGGTATGGGACTTCTTGGTCATCTCCNAGAA-----TAC	966
Qy	259	ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly	278
Db	967	TCTATTTGCACAAGTGCTCACTGTTCTTTTCCGTGTTAAATTGGAGCATTCAGTGTGGA	1026
Qy	279	AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe	298
Db	1027	CAGCATCTCCAAATATTGAAGCTTTGCGCAATGACGAGGAGCAGCTTATGAAGTCTTC	1086
Qy	299	GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu	318
Db	1087	AAATAATTCATTAATGAAGCCAGTAGACAGTCTCAAGAGTGGGCACNAACAGAC	1146
Qy	319	SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer	338
Db	1147	AACATACAAGAAATCTGGAATTTAAGAATATTCTACTTCAGTTACCCATCTCGAAAAAA	1206
Qy	339	IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal	358
Db	1207	GTTCAGATCTTTGAAGGCGCTCAATCTGAAGGTGAAGCGGACGACGCGTGGCCCTGGT	1266
Qy	359	GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro	378
Db	1267	GGCAACAGTGGCTGTGGAAAAACCAACTGTCTCAGCTGATGCAAAAGCTCTACGACCCC	1326
Qy	379	AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr	398
Db	1327	CTAGATGGCATGTGTAGTTCGACGACAGGACATCAGAACCATCAATGTGAGGTATCTG	1386
Qy	399	ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn	418
Db	1387	AGGAGATCATTTGTGTGGTGTGAGTCAGGAACCTGTGCTGTTTGCACCAACCATCGCCGAG	1446
Qy	419	AsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGlu	438
Db	1447	AACATTCGCTATGGCCGGAAGATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGAA	1506
Qy	439	AlaAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu	458
Db	1507	GCCAAATGCCTATGATTCATCAATGAATCCGCCCAACCAATTTGACACCTGGTGGTGAG	1566

Qy	459	LysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal	478
Db	1567	AGAGGGGGCGCAGCTGAGTGGGGGACAGAAACAGAGAATCGCAATTCGCCGGGGCCCTGGCT	1626
Qy	479	ArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys	498
Db	1627	CGCAATCCCAAGATCTTTTGTGGACAGGCCACCTCAGCCCTGGATACAGAAAGTGAA	1686
Qy	499	SerAlaValGlnAlaAlaLeuGlyLysAlaSerLysGlyArgThrThrIleValValAla	518
Db	1687	GCTGTGTTTCAGSCGCACTGGATAAGGCTAGAGAAGCGCGACCACCACTTGTGATAGCT	1746
Qy	519	HisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeu	538
Db	1747	CATCGCTGTGTACCGTTCGTAATCGTACGCTCATGCTGCTTTTGATGGTGGTGCATT	1806
Qy	539	AlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuVal	558
Db	1807	GTGGAGCAGAGAAATCATGATGAGCTCATGAGAGAAAGGCACTTTACTTCAAACTTGTC	1866
Qy	559	MetSerGln-----AspIleLysLysAlaAspGlu	568
Db	1867	ATGACACAGACACAGAGAAATGAAATTGAATTAGGAAATGAAGCTTGTAAATCTCAAGGAT	1926
Qy	569	GlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSer	588
Db	1927	GAATATGATAATTTAGACATGTCTTCAAAAGATTCCAGATCCAGCTTAATAAGAAGA--	1983
Qy	589	VallysSerIleLysSerAspPheIle-----AspLysAlaGluGluSer	603
Db	1984	---AGATCAACTCGCAAAAGCATCTGTGGACACATGACCAGACAGGAAGCTTAGTAGTACC	2040
Qy	604	ThrGlnSerLysGluIleSerLeuProGluValSerLeuLysIleLeuLysLeuAsn	623
Db	2041	AAAGAGGCCCTGGATCAAGATGACTCCAGCTTCCTTTGGCGGATCTCGAAGTGAAT	2100
Qy	624	LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal	643
Db	2101	TCAACTGATGGCTTATTTTGTGGTGTATATCTGTGCCATAATAAATGAGGCGTTA	2160
Qy	644	HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn--AsnAsp	662
Db	2161	CAGCCAGCACTTCCTCGTAATATTTTCAAAAGTTGTAGGGGTTTTTACAAATGTGGCCCC	2220
Qy	663	LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal	682
Db	2221	CCTGAACCCAGCGGCGAGAACAGCAACTGTTTCTGTTGTTCTGTATCTTGGGATC	2280
Qy	683	IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu	702
Db	2281	ATTTCCTTCATTACATTTTTTCTCAGGCTTCACATTTGGCAAGCTGGAGAGATCCTC	2340
Qy	703	ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe	722
Db	2341	ACCAGCGCACTCCGATACATGTTTTCAAATCCATGCTGACACAGGATGTGAGCTGTTT	2400
Qy	723	AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln	742
Db	2401	GATGACCTTAAAAACACCCAGGACATGACCACAGGCTCGCCACAGCATGCTGCTCAA	2460
Qy	743	IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly	762
Db	2461	GTGAAAGGGGCTACAGGGCTAGGCTTGCTGTGATTTTCCAGAACATAGCAAAATCTTGGG	2520
Qy	763	LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle	782
Db	2521	ACAGGAATCATATATCCCTAATCTATGGCTGGCAACTTAACACTTTTACTCTTAGCAATT	2580
Qy	783	AlaProValLeuAlaValThrGlyMetIleGlyThrAlaAlaMetThrGlyPheAlaAsn	802
Db	2581	GTACCCATCATTCGATAGCTGAGGGTGTGAATAATGAATAATGTTGTCTGGCAAGCACTG	2640
Qy	803	LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle	822


```
Db 2641 AAAGATAAGAAAGCAACTAGAAAGGTTCTGGAAGAGATTGCTACGGAAGCAATTTGAAACTTC 2700
Qy 823 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 842
Db 2701 CGCACTGTTCTCTTTGACTCGGGAGCAGAAAGTTGAAACCACTGATGTCGCCAGAGCTTG 2760
Qy 843 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleLeuGlySerCysTyrAlaPhe 862
Db 2761 CAGATACCATACAGAAATGCGATGAAGAAAGCACAGTGTTGGGATCAGCTTCTCCTTC 2820
Qy 863 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 882
Db 2821 ACCAGGCCCATGATGATTTTCTTATGCTGCTGTTTCCGGTTCGGTGCTACTTGGTG 2880
Qy 883 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 902
Db 2881 ACACAACAACCTCATGCTTTTGAATGTTCTGTAGTATTCTCAGCTATGTCTTTGGT 2940
Qy 903 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 922
Db 2941 GCATGCCAGTGGGCGAGTCACTTCCTCGCTCCTGACTATCGGAAGCAACAGTGTCA 3000
Qy 923 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 942
Db 3001 GCATCCACATCATCAGATCATTTGAGAAACCCCGAGATTGCACAGTACAGCACGCAA 3060
Qy 943 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPheTyr 962
Db 3061 GGCTTAAGCCGATATGTTGGGAAGGAAATGTGCAATTTAGTCGAGTCGTGTTCAACTAT 3120
Qy 963 ProCysArgProAspValPheIleLeuArgGlyLysSerLeuSerIleGluArgGlyLys 982
Db 3121 CCACCCGACCCAGCATCCAGTGCTTCAGGGGCTGAGCCCTTGAGGTGAAGAGGGCCAG 3180
Qy 983 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 1002
Db 3181 ACGTGCGCCCTGGTGGGCAGCAGTGGCTGCGGAAGAGCAGACAGTGGTCCAGTGTCTGAG 3240
Qy 1003 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 1022
Db 3241 CGTTCTACACCCCATCGGTGGTGCATCATGCTTCTAGATGCCAAGAAATAAGCAACTG 3300
Qy 1023 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 1042
Db 3301 AATGTCAGTGGCTCCGAGCACAGCTGGGCATTTGTGCCAAGAGCCCATCTCTTTGAC 3360
Qy 1043 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 1062
Db 3361 TGCAGCATCGCAGAGAACATTTGCCCTACGGAGACAACAGCCGGTCTGTCTTTATGAGGAG 3420
Qy 1063 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 1082
Db 3421 ATTGTGAGGCGAGCAAGAGGCGCAACATCCACCGAGTTCATCGACTCGCTACCTGATAAA 3480
Qy 1083 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 1102
Db 3481 TACAAACCAAGATGAGAGACAAGGCACTCAGCTGTCGGGTGGGCGAAGAGCGGCATC 3540
Qy 1103 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 1122
Db 3541 GCCATCGCAGCGCCCTCGTCAGACAGCCTCACATTTTACTTCTGACGAAGCAACATCA 3600
Qy 1123 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 1142
Db 3601 GCTCTGATACAGAAAGTGAAGGTTTGTCCAGAACCGCTGGCAACAGCCAGGAGGC 3660
Qy 1143 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 1162
Db 3661 CGCACTGCAATGTGATCGCTCCCGCTGTCCACCATCCAGAACCGGCACTTGATCGTG 3720
Qy 1163 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 1182
```

```
Db 3721 GTGATTTCAGAACGGCAAGGTCAGGAGCAGCGCACCCACAGCAGCTGCTGGCGCAGAAG 3780
Qy 1183 AspIleTyrPheLysLeuValAsnAlaGlnSer 1193
Db 3781 GGCATCTACTTCTCAATGGTCACTGTGCAGGCT 3813
RESULT 15
US-09-880-107-2299
; Sequence 2299, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2299
; LENGTH: 3924
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 M23234
US-09-880-107-2299
```

```
Alignment Scores:
Pred. No.: 2,38e-306 Length: 3924
Score: 3504.50 Matches: 676
Percent Similarity: 74.43% Conservative: 235
Best Local Similarity: 55.23% Mismatches: 276
Query Match: 58.29% Indels: 37
DB: 10 Gaps: 8
US-09-873-409-6 (1-1195) x US-09-880-107-2299 (1-3924)
Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyValaCysLeuProLeuMetProLeu 20
Db 201 AATGTCCTGGTACCATCATGCGCATAGCTCAGGATCAGGTCTCCCTCATGATGATA 260
Qy 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
Db 261 GTATTTGGAGAGATGACTGACAAATTTGTTGAT-----ACTGCAGGAAACTTC 308
Qy 41 SerPhe-----PheArgLeuThrLeu----- 47
Db 309 TCCTTTCCAGTGAACATTTTCCTTGCTGCTGCTAAATCCAGGCAAAATTCGGAAGAGAA 368
Qy 48 -----TyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db 369 ATGACTAGATATGCATATTACTACTCAGGATTTGGGTGCTGGAGTCTTGTGCTGCTAT 428
Qy 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db 429 ATACAAGTTTCAATTTTGGACTTTTGGCAGCTGGTTCGACAGATCAGGAAATTTAGGCAGAAG 488
Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db 489 TTTTTCATGCTATTCTACGACAGGAAATAGGATGTTTGACATCAATGACACCACTGAA 548
Qy 102 LeuAsnThrArgMetThr--AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db 549 CTCATATCCGGGTACACATGACATCTCCAAATCAGTGAAGAAATTTGGTGCACAGGTT 608
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
```


Db 609 GGAATGTTCTTTCAAGCAGTAGCCACGTTTTTTTGGCAGGATTCATAGTGGGATTCATCAGA 668
Qy 141 GlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db 669 GGATGAAGCTCACCTTGTGATAAATGGCCATCAGCCCTATTCTAGGACTCTCTGCGAGCC 728
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db 729 GTTGGCCAAAGATCTCTCGCATTTAGTGACAAAGAACTAGCTGCTTATGCAAAAGCA 788
Qy 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db 789 GGGCCGTCGACAGAGAGGCTCTGGGGCCATCAGACTGTGTAGCTTTTCGGGGCCAG 848
Qy 201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
Db 849 AACAAAGAGCTGGAAGGATCAGAAACATTTAGAAAAATGCCAAAGAGATTGGAATTTAA 908
Qy 221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
Db 909 AAAGCTATTTCAGCAAAACATTTCCATGGGTATTGCCCTTCCTGTTAATATATATGCATCAT 968
Qy 241 GlyLeuAlaPheTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
Db 969 GCACCTGGCCTTCGGTATGGATCCACTCTAGTCATATCAAAAGAA-----TATACTATT 1022
Qy 261 GlyThrValLeuAlaValPheSerValIleHisSerTyrCysIleGlyAlaAla 280
Db 1023 GGAATGCAATGACAGTTTTTTTCAATCCTAATTTGGAGCTTTTCAGTGTGCCAGGCT 1082
Qy 281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
Db 1083 GCCCATGTATTGATGCTTTTGGCAATGCAAGAGGAGCAGCATATGTGATCTTTTGATATT 1142
Qy 301 IleAspLysIleProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
Db 1143 ATTGATTAATAATCCTAAATTTGACGTTTTTCAGAGAGGAGCAACACAGACGATC 1202
Qy 321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
Db 1203 AAAGGAATTTGGAGTTCAATGATGTTTCACTTTTACCTTCTCGAGCTAACGTCAG 1262
Qy 341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
Db 1263 ATCTTGAAGGGCTCAACCTGAAGGTGCAGAGTGGCGACAGCGTGGCCCTGGTTGGAAGT 1322
Qy 361 AsnGlySerGlyLysSerThrValValGlnLeuGlnArgLeuTyrAspProAspAsp 380
Db 1323 AGTGGCTGTGGGAAGAGCACACGGTCCAGCTGATACAGAGGCTCTATGACCCCTGATGAG 1382
Qy 381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
Db 1383 GGCACATTTAAACATTGATGGCGAGGATTTAGGAATTTAATGTAACTATCTGAGGGAA 1442
Qy 401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
Db 1443 ATCATTTGGTGTGAGTCAGAGCGGTGCTGTTTCCACCACCAATTTGCTGAAATATT 1502
Qy 421 LysTyrGlyArgAspValThrAspGluLysGluMetGluArgAlaAlaArgGluAlaAsn 440
Db 1503 TGTATTGGCCGTGGAATGTAAACCATGGATAGATAAAGAGCTCTCAAAAGAGGCCAAC 1562
Qy 441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
Db 1563 GCCTATGAGTTTATCATGAAATTTACCACAGAAATTTGACACCCCTGGTTGGAGAGAGGG 1622
Qy 461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db 1623 GCCCAGCTGATGGTGGCAGAGAGAGGATGCGCATTCGACGTGCCCTGGTTTCGCAAC 1682
Qy 481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db 1683 CCCAAGATCCTTCTGCTGATGAGGCCACGTCAGCATTTGGACACAGAAAAGTGAAGCTGAG 1742

Qy 501 ValGlnAlaAlaLeuGluLysAlaSerTysGlyArgThrThrIleValValAlaHisArg 520
Db 1743 GTACAGGCAGCTCTGGATAAGCCAGAGAAGCGCGACCATTTGTGTAGTACACACCGA 1802
Qy 521 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 540
Db 1803 CTGTCTACGGTCCGAAATGCAGATGTCATCGCTGGGTTTTGAGGATGGAGTAATTGGGAG 1862
Qy 541 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuVal----- 558
Db 1863 CRAAGNAGCCACAGCACTGATGAAGAAGAGGGGTGCTACTTCAAACTTGTCAACATG 1922
Qy 559 -----MetSerGlnAspIleLysLysAlaAspGluGlnMetGlu 571
Db 1923 CAGACATCAGGAAGCCAGATCCAGTCAGAAAGATTTGAACTAAATGATGAAAAGCTGCC 1982
Qy 572 SerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSer 591
Db 1983 ACTAGAATGGCCCAAAATGGCTGGAAATCTCGCCTATTATTAGGCATTTCTACTCAGAAAAAC 2042
Qy 592 IleLys---SerAspPheIleAspLysAla---GluGluSerThrGlnSerLysGluIle 609
Db 2043 CTTAAAAATTCAAATGTGTGAGAGAGCCTTGATGTGGAACCGATGGACTTGAGCA 2102
Qy 610 SerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTyrProPhe 629
Db 2103 AATGTGCCACACAGTGTCTTCTGAAGGTCTCTGAACTGAATAAACAGAAATGGCCCTAC 2162
Qy 630 ValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIle 649
Db 2163 TTTTGTGTGGGACAGATATGTGCCATTGGCAATGGGGGCTTCAGCCGCAATTTTCAGTC 2222
Qy 650 IlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThrLeuLysHisAsp 669
Db 2223 ATATTCTCAGACATCATGCGAATTTTGGACCAGCGGATGTCAGTGAAGCAGCAGAAAG 2282
Qy 670 AlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPhe 689
Db 2283 TGCAACATATTCTCTTGATTCTTATTCTTCTGGGAATTTATTCTTTTATTCTTCTTC 2342
Qy 690 MetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeu 709
Db 2343 CTTTCAGGGTTTCACGTTTGGGAAAGCTGGCGAGATCCTCACCAGAAAGACTGCGGTCAAT 2402
Qy 710 AlaPheLysAlaMetLysTyrGlnAspIleAlaTyrPheAspGluLysGluAsnSerThr 729
Db 2403 GCTTTTAAAGCAATGCTAAGCAGGACATGAGCTGTTTGTGATGACCATATAAAACAGTACT 2462
Qy 730 GlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySer 749
Db 2463 GGTGCACATTCTACAAGACTTGCCACAGATGCTGCCAAGTCCAAGGAGCCACAGNAACC 2522
Qy 750 ArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPhe 769
Db 2523 AGGTTGGCTTTAATGGCAGAAATATAGCTAACTTGGAACTGGTATATCATATCATTT 2582
Qy 770 IleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThr 789
Db 2583 ATCTACGGTGGCAGTTAACCCCTATTGCTATTAGCAGTTGTTTCCAAATTTATGCTGTGCTCA 2642
Qy 790 GlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnLeuLys 809
Db 2643 GGAATGTTGAAATGAAATTTGGCTGGAATGTCAAAAGAGATATAAAAGAACTGGAA 2702
Qy 810 HisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThr 829
Db 2703 GCTGTGGAAGATTTGCAACAGAGCAATAGAAAATATTAGGACAGTTGTGTCTTTGACC 2762
Qy 830 ArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThr 849
Db 2763 CAGGAAGAAAAATTTGAAATCAATGTATGTTGAAAAAATTTGTGAGCACTTTACAGGAATTC 2822

Search completed: March 31, 2003, 16:13:21
Job time : 686.971 secs

```
QY 850 SerLysLeuAlaGlnIlelleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPhe 869
Db 2823 GTCCAGAGGCACACATCTATGGAATTACTTTAGTATCTCACAAGCATTTATGATATTT 2882
QY 870 AlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrPro 889
Db 2883 TCCTATGCGGGTGTGTTTCGATTGTGTGCATATCTCATTTGTAATGCACATATGCGCTTC 2942
QY 890 GluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThr 909
Db 2943 AGAGATGTTTATCTGGTGTGTTTCTGCAATTGTATTTGGTGCAGTGGCTCTAGGACATGCC 3002
QY 910 LeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeu 929
Db 3003 AGTTTCATTTGCTCCAGACTATGCTAAAGCTAAGCTAGCTGTGCAGCCCATTTATTCATGCTG 3062
QY 930 LeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCys 949
Db 3063 TTTGAAAGACACACTCTGATTGACAGCTACAGTGAAGAGGGGCTGAAGCCTGATAAATTT 3122
QY 950 GluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPhe 969
Db 3123 GAAGGAAATATAACATTTAATGAAGTCGTCTCACTATCCACCCGAGCAACGTCGCCA 3182
QY 970 IleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySer 989
Db 3183 GTGCTTCAGGGGCTGAGCTTGAGGTGAAGAAAGGCCAGACACTAGCCCTGGTGGGCAGC 3242
QY 990 SerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnAlaArgLeuTyrAspProValGln 1009
Db 3243 AGTGGCTGTGGGAAGACACGGTGTCCAGCTCCTGGAGCGGTCTACGACCCCTTGGCG 3302
QY 1010 GlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTyrLeuArgSer 1029
Db 3303 GGGACAGTGTCTTCGATGTGTCAAGAGCAAGAACTCAATGTCCAGTGGCTCAGAGCT 3362
QY 1030 GlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIle 1049
Db 3363 CAATCGAATCGTGTCTCAGGAGCTATCTATTGTGACTGCGACATTGCCGAGATATT 3422
QY 1050 AlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAla 1069
Db 3423 GCCTATGGAGACAAACAGCCGGTGTATCACAGATGAAATTTGTAGTGCAGGCCAAAGCT 3482
QY 1070 AlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeu 1089
Db 3483 GCCAACATACATCCTTTTCATCGAGACGTTACCCCAAAATATGAAACAGAGTGGGAGAT 3542
QY 1090 LysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeu 1109
Db 3543 AAGGGACTCAGCTTCAGGAGGTCAAAAACAGAGGATTGCTATTGCCCGAGCCCTCATC 3602
QY 1110 GlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGlu 1129
Db 3603 AGACAACCTCAAATCCTCTGTTGGATGAAGCTACATCAGCTCTGGATACTGAAAGTGAA 3662
QY 1130 LysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThr 1149
Db 3663 AAGGTTGTCAAAGAAGCCCTCGCAAGCCGAGAGAGCCGACCTGCAATTTGATTTGCT 3722
QY 1150 HisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIle 1169
Db 3723 CACCGCTGTCCACCATCCAGATGCAGACTTAATAGTGTGTTTCAGATGGGAGATC 3782
QY 1170 LysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuVal 1189
Db 3783 AAGGAGCATGGCAGCATCAGACGCTGCTGGCACAGAAAGGCATCTATTTTCAATGGTC 3842
QY 1190 AsnAlaGlnSer 1193
Db 3843 AGTGTCAGGCT 3854
```


GenCore version 5.1.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 30, 2003, 01:24:25 ; Search time 16155.1 Seconds
(without alignments)
2152.747 Million cell updates/sec

Title: US-09-873-409-6
Perfect score: 6012
Sequence: 1 MILGILASLVNGACPLMPL.....QELLNRNDIYFKLVNAQSVQ 1195

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DEV=xlh
-O=/cgn2.1/USPTO.spool/US09873409/runat_27032003.115418.19228/app_query_fasta_1.7544
-DB=GenEmbl -QFMT=fastap -SURFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09873409 @CGN 1.1.54887 @runat_27032003.115418.19228 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_WAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.ma.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*

29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	6012	100.0	3621	6	AX339032 Sequence
2	5988.5	99.6	3702	6	AX339031 Sequence
3	5760.5	95.8	3699	6	AX478104 Sequence
4	5159.5	85.8	3177	6	AX339030 Sequence
5	4561.5	75.9	2856	6	AX339028 Sequence
6	3574	59.4	4296	10	CRUPGPI
7	3573	59.4	4304	10	CRUPGPI165
8	3557	59.2	4279	6	AX105082
9	3556.5	59.2	4195	6	AX108656 Sequence
10	3555.5	59.1	4927	10	AF257746
11	3554	59.1	3860	6	AX322787 Sequence
12	3554	59.1	4279	6	AX105078 Sequence
13	3554	59.1	4279	6	AX105080 Sequence
14	3554	59.1	4646	6	AX391099 Sequence
15	3552	59.1	4045	12	AF269224
16	3551	59.1	4378	6	E02326
17	3549.5	59.0	3858	4	OAU78609
18	3549	59.0	4279	6	AX105057 Sequence
19	3549	59.0	4317	4	AF045016
20	3549	59.0	4317	6	AX105059 Sequence
21	3548	59.0	4186	6	AX108654 Sequence
22	3545	59.0	3860	6	AX322789 Sequence
23	3545	59.0	4646	6	AX336420 Sequence
24	3545	59.0	4646	6	AX336708 Sequence
25	3545	59.0	4646	6	I49610
26	3545	59.0	4646	9	HUMMDR1
27	3545	59.0	4669	6	I08557
28	3545	59.0	6505	6	AR028671 Sequence
29	3545	59.0	8630	6	AX012320 Sequence
30	3545	59.0	8630	6	AX012321 Sequence
31	3545	59.0	9318	6	AX028672 Sequence
32	3543.5	58.9	4264	6	AR051647 Sequence
33	3543.5	58.9	4264	6	AR051650 Sequence
34	3541.5	58.9	4323	10	AF286167
35	3537.5	58.8	3840	6	AX481416
36	3537.5	58.8	3988	6	AX024454 Sequence
37	3537.5	58.8	4192	9	AF016535
38	3529	58.7	4280	10	CRUPGPII
39	3527	58.7	4669	6	AR091275 Sequence
40	3527	58.7	4669	6	AR203322 Sequence
41	3526.5	58.7	4189	6	AX322791 Sequence
42	3526.5	58.7	4298	10	MUSMDR
43	3523	58.6	3934	4	CFA419568
44	3522.5	58.6	3888	10	AY082609
45	3522.5	58.6	4788	6	AX322793 Sequence

ALIGNMENTS

RESULT 1

Db	1714	CAGGATATTAAAAAGCCTGATGACAGATGGAGTCAATGACATATTTCTACTGAAAGAAG	1773
Qy	581	ThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAla	600
Db	1774	ACCACTCACTTCTCTGTGCACCTCTGGAAGACATCAAGTCAGACTTCATTCGACAGGCT	1833
Qy	601	GluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeu	620
Db	1834	GAGGAATCCACCCAACTTAAAGAGATAAGCTTCTCTGAAGTCTCTCTATTAAAAAATTTTA	1893
Qy	621	LysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsn	640
Db	1894	AAGTTAAACAAGCCTGAATGGCCCTTTGTGGTTCCTGGGGACATTTGGCTCTGTGTTAAAT	1953
Qy	641	GlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn	660
Db	1954	GGAACTGTTTCATCCAGTATTTTCCATCATCTTTTGCAAAAATTAATCAACGTGTTGGAAAT	2013
Qy	661	AsnAspLysThrThrLeuLysHisAspAlaGluIleIleTyrSerMetIlePheValIleLeu	680
Db	2014	AATGATAAAACCAACATTTAAAGCATGATGCAGAAAATTTATTCATGATATTCGTCAATTTG	2073
Qy	681	GlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGlu	700
Db	2074	GGTGTTATTTGCTTTGTGCAGTTATTTTCATGCAGGAGATATTTTACGGCAGACGAGGGAA	2133
Qy	701	IleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAla	720
Db	2134	ATTTTAAACGATGAGATTAGACACTTTGGCCCTTCAAGCCATGTTATATCAGGATATTGCC	2193
Qy	721	TrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIle	740
Db	2194	TGGTTTGATGAAAAGGAAAACAGCAGCAGAGGCTTGACACAATATTAGCCATAGATATA	2253
Qy	741	AlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsn	760
Db	2254	GCACAAATTCACAGGAGCAACAGGTTCCAGGATVGGCGTCTTAAACACAAAATGCAACTAAC	2313
Qy	761	MetGlyLeuSerValIleIleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeu	780
Db	2314	ATGGGACTTTTCAGTTATCATTTTCCCTTTATATGATGAGGAGATGACATTCCTGATTCGT	2373
Qy	781	SerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPhe	800
Db	2374	AGTATTGCTCCAGTACTTGGCGTGACAGAGATGATTGAAACCGCAGCAATGACTGGATTT	2433
Qy	801	AlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGlu	820
Db	2434	GCCAAACAAAGATAAGCAAGAACTTAAAGCATGCTGGAAGATAGCAACTGAAGCTTTCGAG	2493
Qy	821	AsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGlu	840
Db	2494	AATATACGCTACTATAGTGTTCATTTAAACAGGGGAAAAGCCCTTCAGCAAAATGATGAAGAG	2553
Qy	841	MetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyr	860
Db	2554	ATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCACAGATTTATTGGAAAGCTGTAT	2613
Qy	861	AlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyr	880
Db	2614	GCAATTCAGCCATGCTTTATATATTTTGGCTTCATGCAGCAGAGGGTTCGATTTGGAGCCTAT	2673
Qy	881	LeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAla	900
Db	2674	TTAATTCACGCTGCAGCAATGACCCACAGAGGGCATGTTTCATAGTTTTTACTGCAATTTGCA	2733
Qy	901	TyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLys	920
Db	2734	TATGGAGCTATGGCCATCGGAAAACGCTCGTTTTGGCTTCCTGAAATATATCCAAAGCCAAA	2793
Qy	921	SerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSer	940

Db	2794	TCGGGGGCTCGCATCTCTGTTGGCTTGTGGAAAGAAACCAATATATAGACAGCGCGAGT	2853
Qy	941	GlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhe	960
Db	2854	CBAGAAGGGAAAAGCCAGACACATGTGAAGGGAATTTAGAGTTTCGAGAAGTCTCTTTC	2913
Qy	961	PheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArg	980
Db	2914	TTCTATCATGTGCGCCAGATGTTTTCATCTCCGTGGCTTATCCCTCAGTATTGAGCGA	2973
Qy	981	GlyLysThrValAlaPheValIcylSerSerGlyCysGlyLysSerThrSerValGlnLeu	1000
Db	2974	GGAAAGACAGTAGACATTTGTGGGAGCAGCGCTGTGGGAANAAGCATCTTCGTTCAACTT	3033
Qy	1001	LeuGlnArgLeuTyrAspProValGlnGlyValLeuPheAspGlyValAspAlaLys	1020
Db	3034	CTGCAGAGACTTTATGACCCCGTCAAGAGCAAGTGTGTTTGTATGGTGTGGATGCAAAA	3093
Qy	1021	GluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeu	1040
Db	3094	GAATTGAATGTACAGTGGCTCGGTTCCCAATAGCAATTCGTTCTCAAGACGCTGTGCTC	3153
Qy	1041	PheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeu	1060
Db	3154	TTCAACTCGAGCATTTGCTGAGAAATCCGCTATGTGTGACACAGCGGTGTGGTGCATT	3213
Qy	1061	AspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuPro	1080
Db	3214	GATGAGATCAAGAAGCGCGCAATATGCACAAATATCCATCTTTTATTGAAGGCTCCCT	3273
Qy	1081	GluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGln	1100
Db	3274	GAGAAATACAAACACACAAAGTTGGACTGAAGAGGAGCACAGCTTCTTGGCGGCGCAAA	3333
Qy	1101	ArgLeuAlaIleAlaArgAlaLeuGlnLysProLysIleLeuLeuLeuAspGluAla	1120
Db	3334	AGACTAGCTATTGCAAGGCTCTTCTCCAAAAACCCAAATTTTATTGTTGGATGAGGCC	3393
Qy	1121	ThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArg	1140
Db	3394	ACTTCAGCCCTCGATATGACAGTGAGAGGTGGTTTCAGCATGCCCTTGATANAAGCCAGG	3453
Qy	1141	ThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeu	1160
Db	3454	ACGGGAAGGACATGCCTAGTGCTCACTCACAGGCTCTCGCAATTCAGACGCAGATTG	3513
Qy	1161	IleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArg	1180
Db	3514	ATATGTGGTTCTGCACAAATGGAAAGATAAAGGAACAAGAATCATCAAGAGCTCTGTAGA	3573
Qy	1181	AsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln	1195
Db	3574	AATCAGACATATATTTTAAGTTAGTAGTGAATGCACAGTCAGTGCAG	3618
RESULT 2			
AX339031			
LOCUS	AX339031	3702 bp	DNA linear PAT 09-JAN-2002
DEFINITION	Sequence 13 from Patent WO0194400.		
ACCESSION	AX339031		
VERSION	AX339031.1	GI:18129123	
KEYWORDS	human.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1.		
JOURNAL	Frank, M.H. and Sayegh, M.H.		
FEATURES	A gene encoding a multidrug resistance human p-glycoprotein		
	homologue on chromosome 7p15-21 and uses thereof		
	Patent: WO 0194400-A 13 13-DEC-2001.		
	THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)		
	Location/Qualifiers		
	1..3702		

/organism="Homo sapiens" /db_xref="taxon:9606"									
BASE COUNT	1098 a	708 c	856 g	1039 t	1 others				
ORIGIN									
Alignment Scores:									
Pred. No.:	0	Length:	3702						
Score:	5988.50	Matches:	1195						
Percent Similarity:	97.79%	Conservative:	0						
Best Local Similarity:	97.79%	Mismatches:	0						
Query Match:	99.61%	Indels:	27						
DB:	6	Gaps:	1						
US-09-873-409-6 (1-1195) x AX339031 (1-3702)									
QY	1	MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu	20						
DB	34	ATGATCCTGGGTATACTGGCATCACTGGTCAATGGAGCCTGCCCTTCCTTTAAATGCCACTG	93						
QY	21	ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr	40						
DB	94	GTTTTAGGAGAAATGAGTGATAACCTTATTAGTGGATGCTAGTCCAAACTAACACATAC	153						
QY	41	SerPhePheArgLeuThrIleuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGly	60						
DB	154	TCITTTCTTCAGGTTGACCTGTATTATGTTGGAATAGGTGTTGCTGCCCTTGATTTTGGT	213						
QY	61	TyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLys	80						
DB	214	TACATACAGATTTCTCTGTGGATTATAACTGCAGCACAGACCAAGAGGATTCGAAAA	273						
QY	81	GlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGly	100						
DB	274	CAGTTTTTTCATTTCAGTTTTGGCACAGGACATCGCTGGTTTGATAGCTGTGACATCGGT	333						
QY	101	GluLeuAsnThrArgMetThrAspIleAspLysIleSerAspGlyIleGlyAspLysIle	120						
DB	334	GAACCTTAACACTCGCATGACAGACATTTGACAAATCAGTCAGTGTATTGGAGATAAGATT	393						
QY	121	AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys	140						
DB	394	GCTCTGTGTTTCAAAAATGCTACTTTTTTCGATTGGCCCTGGCAGTTGTTTGGTGAAG	453						
QY	141	GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla	160						
DB	454	GGCTGGAAACTCACCTAGTAGCTCTATCCAGCTCTCTCTTATATAGCTTCAGGGCA	513						
QY	161	AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla	180						
DB	514	GCATGTTCTAGGATGGTCATCTCATTGACCCAGTAAGGAATTAAGTGCTTATTCCAAAGCT	573						
QY	181	GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln	200						
DB	574	GGGGCTGTGGCAGAAGAAGTCTTTGTCTATCAATCCGAACAGTCATAGCCTTTTAGGGCCACG	633						
QY	201	GluLysGluLeuGln-----	205						
DB	634	GAGAAAGAACTTCAAGGCTTTTCCTTTTAAATATAACAGATATGCTTGGTTTATTTT	693						
QY	206	-----ArgTyrThrGlnAsnLeuLysAsp	213						
DB	694	CCCCAGTGGCTACTAAGTTGTGTTCTGTTNTTGTAAAGGTATACACAGAATCTCAAAGAT	753						
QY	214	AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr	233						
DB	754	GCAAGAGATTTTGGCATAAAAAGGACTATAGCTTCAAAAGTGTCTCTTGGTGTCTGTGTAC	813						
QY	234	PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn	253						
DB	814	TTCTTTATGAATGGAACCTATGGACTTGCCTTTTGGTATGGAACCTCTCTTGATTCCTTAAT	873						
QY	254	GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer	273						

DB	874	GGAGAACTCGGATATACCATCGGAGCTGTTCTTGTGCTGTTTCTTTAGTGTAAATCCATAGC	933
QY	274	SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaAArgGlyAla	293
DB	934	AGTTATTGCATTGGAGCAGCAGTCCCTCACTTTTGAACCTTCGCAATAGCCCGAGAGCT	993
QY	294	AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla	313
DB	994	GCCTTTTCATATTTCCAGGTTATTGATAGAAACCCAGCATATAGATAACTTTTCCACAGCT	1053
QY	314	GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr	333
DB	1054	GGATATAAAACCTGAATCCATAGAGGAACCTGTGGAATTTAAATAATGTTCTTTCAATAT	1113
QY	334	ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu	353
DB	1114	CCATCAAGACCATCTATCAAGATTCTGAAGGCTCGAACTCTCAGAAATTAAGTCTGGAGAG	1173
QY	354	ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln	373
DB	1174	ACAGTCGCCTTGGTCGCTCTCAATGGCAGTGGGAAGAGTAGTGTCCAGCTTCTGCAG	1233
QY	374	ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu	393
DB	1234	AGTTTATATGATCCGATGATGCTTTTATCATGGTGGATGAGAATGACATCAGAGCTTTA	1293
QY	394	AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly	413
DB	1294	AATGTGGGCAATTATCGAGACCAATTTGGAGTGGTTAGTCAAGAGCTGTTTGTTCGGG	1353
QY	414	ThrThrIleSerAsnAsnIleLysTyrGlyArgAspValThrAspGluGluMetGlu	433
DB	1354	ACCACCATCAGTAACAAATATCAAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAG	1413
QY	434	ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProLeuLysPheAsn	453
DB	1414	AGAGCAGCAAGGGAAGCAAAATGCTATGATTTTATCATGGAGTTTCTTAATAATTAAT	1473
QY	454	ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle	473
DB	1474	ACATTTGGTGGGAAAGAGAGCTCAAAATGAGTGGAGGGGAGAAACAGAGATCGCAANT	1533
QY	474	AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu	493
DB	1534	GCTCGTGCCTTAGTTCGAAACCCCAAGATTCTGATTTTAGATGAGGCTAGCTCTGCCCTG	1593
QY	494	AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThr	513
DB	1594	GATTCAGAAAGCAAGTCAGCTGTTCAAGCTGCACCTGGAGAAGCGGAGCAAGGTCGGACT	1653
QY	514	ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu	533
DB	1654	ACAATCGTGTGATACACCGACTTTCTATTTCGAAGTGCAGATTTGATGTGACCCCTA	1713
QY	534	LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu	553
DB	1714	AAGGATGGAATGCTGGCGGAGAAAGGAGCACATGCTGAATAATGGCAAAACAGAGTCTA	1773
QY	554	TyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMet	573
DB	1774	TATATTTCATCTGTGATGTACAGGATATTTAAATAAGCTGATGAACAGATGAGTCAATG	1833
QY	574	ThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLys	593
DB	1834	ACATATTCTACTGAAAGAAAGACCAACTCTCTCTCTGCACTCTGTGAAGAGCATCAAG	1893
QY	594	SerAspPheIleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGlu	613
DB	1894	TCAGACTTCTATTCACAAAGCTGAGGAATCCACCAATCTAACAGAGATAAGTCTCTCTGAA	1953
QY	614	ValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGly	633
DB	1954	GTCTCTCTATTAATAAATTTTAAAGTAAAAACAGCCTGAATGGCCTTTTGTGGTCTCGGG	2013

QY 634 ThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLys 653
Db 2014 ACATTGGCTTCTGTTCTAAATGGAACGTTCATCCAGTATTTTCCATCATCTTTGCAAAA 2073
QY 654 IleIleThrMetPheGlyAsnAsnAspLysThrLeuLysHisAspAlaGluIleTyr 673
Db 2074 ATTATACCATGATGTTTGGAAATATATGATAAACCACATTAAGCATATGACAGAAATTTAT 2133
QY 674 SerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeu 693
Db 2134 TCATGATATATTCGTATTTGGTGATATTTGCTTTGTCAGTTATTTTCATGACGGGATTA 2193
QY 694 PheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheTysAla 713
Db 2194 TTTTACGGCAGACGCGGAAATTTTAAACGATGAGATTAAACACACTTGGCCTTCAAGCC 2253
QY 714 MetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThr 733
Db 2254 ATGTTATATACGAGATATTCCTGGTGTGATGAAAGGAAACACACAGAGGCTTGACA 2313
QY 734 ThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyVal 753
Db 2314 ACATATTAGCCATAGATATAGCAAAATTCAGGAGCAACAGGTTCCAGGATTGGCGTC 2373
QY 754 LeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyr 773
Db 2374 TTAACACAAAATGCAACTAACATGGGACCTTCAGTTATCATTTCTTTATATATGATGG 2433
QY 774 GluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGlu 793
Db 2434 GAGATGACATTTCTGATTTCTGATTTGCTCCAGTACTTTCGCGTGACAGGAATGATTGAA 2493
QY 794 ThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLys 813
Db 2494 ACCGACGCAATGACTGGATTTGCCACAAAGATAAGCAAGAACTTAAGCATGCTGGAAG 2553
QY 814 IleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAla 833
Db 2554 ATAGCAACTGAACTTTGAGAAATATACGTACTATATAGTGTCTTAACAGGGAAGAAGCC 2613
QY 834 PheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAla 853
Db 2614 TTGAGCAAAATGATGAAGAGATGCTTTCAGACTCAACACAGAAATACCTCGAAGAAGCA 2673
QY 854 GlnIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaValAlaAla 873
Db 2674 CAGATTATTGGAAGCTGTATGCATTCAGCCATGCGCTTTATATATTTTGCCTATGCAGCA 2733
QY 874 GlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPhe 893
Db 2734 GGGTTTCGATTGGAGCCTATTAAATCAAGCTGGAGCAATGACCCAGAGGCGCATGTTTC 2793
QY 894 IleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAla 913
Db 2794 ATAGTTTTTACTGCAATTCATATGAGCTATGCGCATGCGCAATCGAAAAACGCTCGTTTGGCT 2853
QY 914 ProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuGluLysLys 933
Db 2854 CCTGAATATTCCAAAGCCAAATCGGGGCTGCGCATCTGTTTCCCTTGTGGAAGAA 2913
QY 934 ProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeu 953
Db 2914 CCNAATATAGACAGCCGCGAGTCGAAGAGGGAAGAAAGCCAGACATGTGAGGGAATTTA 2973
QY 954 GluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGly 973
Db 2974 GAGTTTCGAGAGACTCTCTTCTTCTATCCATGTCGCCAGATGTTTTCATCTCCCTGGC 3033
QY 974 LeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGly 993
Db 3034 TTATCCCTCAGTATTGAGCGAGGAAAGACAGTAGCATTTTGTGGGAGCAGCGCTGTGGG 3093

QY 994 LysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeu 1013
Db 3094 AAAAGCACTTCTGTTCAACTTCTGAGAGACTTTATGACCCCGTGCAGGACAAAGTGCTG 3153
QY 1014 PheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIle 1033
Db 3154 TTTGATGGTGTGGATGCAAAAGAAATTGAATGATACAGTGGCTCCGTTCCCAAAATAGCAATC 3213
QY 1034 ValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAsp 1053
Db 3214 GTTCTCAAGAGCCTGTGCTTCAACTGCAGCATTTGCTGAGAACATCGCTATGTGTGAC 3273
QY 1054 AsnSerArgValValProLeuAspGluIleLysGluAlaAsnAlaAlaIleHis 1073
Db 3274 AACAGCCGTGTGGTGCATTTAGATGAGATCAAGAAGCGCAAAATGCAGCAATATATCCAT 3333
QY 1074 SerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGln 1093
Db 3334 TCTTTTATTGAAGGCTCTCCCTGAGAAATACAAACAAGTTGGACTGAAGAGGACACAG 3393
QY 1094 LeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLys 1113
Db 3394 CTTTCTGGCGCCAGAAACAAAGACTAGCTATTGCAAGGGCTCTTCTCCAAAAACCCAAA 3453
QY 1114 IleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGln 1133
Db 3454 ATTTTATTGTTGATGAGGCCACTTTCAGCCCTCGAATAATGACAGTGCAGAGGTGGTTTCAG 3513
QY 1134 HisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSer 1153
Db 3514 CATGCCCTTGATAAAGCCAGGACGGGAAGACATGCTTAGTGGTCACTCACAGGCTCTCT 3573
QY 1154 AlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGly 1173
Db 3574 GCATTCAGAACGCAGATTTGATGTTGTTCTGCACATGGAAGATTAAGGACAAAGGA 3633
QY 1174 ThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1193
Db 3634 ACTCATCAAGAGCTCTCGAATAATCGAGACATATATTTTAAGTTAGTGAATGCACAGTCA 3693
QY 1194 ValGln 1195
Db 3694 GTGCAG 3699
RESULT 3
LOCUS AX478104 3699 bp DNA linear PAT 12-AUG-2002
DEFINITION Sequence 34 from Patent WO0240541.
ACCESSION AX478104
VERSION AX478104.1 GI:22217064
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Tang, Y.T., Yue, H., Nguyen, D.B., Hafalia, A.J., Elliott, V.S., Lu, Y.,
Walia, N.K., Yao, M.G., Baughn, M.R., Gandhi, A.R., Ding, L.,
Sanjanwala, M., Ramkumar, J., Arvizu, C., Gietzen, K.J., Lal, P.G.,
Azimzai, Y., Khan, F.A., Thangavelu, K., Thornton, M., Lu, D.A.,
Tribouley, C.M., Warren, B.A., Ison, C.H., Das, D., Raumann, B.E.,
Policky, J.L. and Kearney, L.
TITLE Transporters and ion channels
JOURNAL Patent: WO 0240541-A 34 23-MAY-2002;
FEATURES Location/Qualifiers
source 1. 3699
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="incyte ID No: 7472030CB1"
BASE COUNT 1116 a 707 c 860 g 1016 t
ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 3699
Score: 5760.50 Matches: 1156
Percent Similarity: 95.87% Conservative: 4
Best Local Similarity: 95.54% Mismatches: 9
Query Match: 95.82% Indels: 41
DB: 6 Gaps: 3

US-09-873-409-6 (1-1195) x AX478104 (1-3699)

QY 1 MetIleuGlylleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
DB 145 ATGATCTGGTATCTACTGACATCACTGTTCAATGGAGCCTGCTTCTTTAAATGCCACTG 204
QY 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
DB 205 TGTATAGGAGAAATAGTGTAAACCTTATTAGTGGATGTCTAGTCCACACTAACACAACA 264
QY 41 SerPhePheArg-----LeuThr 46
DB 265 AATTATCAGAAGCTACTCAGTCTCAAGAGAGCTGAATGAATATAGACTCTGTTGACC 324
QY 47 LeuTyrTyrValGlylleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeu 66
DB 325 CTGTATTATGTTGGAATAGGTGTTGCTGCTTGTATTTTGGTTACATACAGATTTCTCTG 384
QY 67 TrpIleIleThrAlaAlaArgGlnThrLysArgIleArgIleGlyGlnPhePheHisSerVal 86
DB 385 TGAATTATAACTGAGCAGCACAGACCAAGAGGATTCGAAACAGTTTTTTTCATTGAGTT 444
QY 87 LeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGluLeuAsnThrArgMet 106
DB 445 TTGGCAGAGACATCGGCTGTTTGTATAGTGTGACATCGTGAACCTTAACACTCGCATG 504
QY 107 Thr---AspIleAspIlylleSerAspGlyIleGlyAspIlylleAlaLeuLeuPheGln 125
DB 505 ACAGATGACATTCACAAAATCAGTGATGTTATGGAGATAAGATTGCTCTGTTGTTTCAA 564
QY 126 AsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValIysGlyTyrPlyLeuThr 145
DB 565 AACATGTCATCTTTTCGATTGGCCTGGCAGTTGGTTTGTGAGGGCTGGAACCTCACC 624
QY 146 LeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAlaCysSerArgMet 165
DB 625 CTAGTGACTCTATCCAGCTCTCTCTTATATGCTTCAGCGGAGCATGTTCTAGGATG 684
QY 166 ValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGlu 185
DB 685 GTCATCTCTATTGACCAAGGAATTAAGTGCCTATTCCAAAGCTGGGGCTGTGGCAGAA 744
QY 186 GluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGln 205
DB 745 GAAGTCTGTTCATCAATCCGAACAGTCATAGCCTTTTAGGCCCCAGGAGAAAGAACTTCAA 804
QY 206 ArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSer 225
DB 805 AGGTATACAGAAATCTCAAGAATGCAAGATTTTGGCATTAAGAAGCACTATAGCTTCA 864
QY 226 LysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTyr 245
DB 865 AAAGTGTCTCTGTGTGTGTCTCTTTTGAATGGAACCTTATGGACTTGTCTTTTGG 924
QY 246 TyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAla 265
DB 925 TATGGAAACCTCTTGATTTCTTAATGGAGAACCTTGGATATACCTCGGGACTGTTCTGCT 984
QY 266 ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGlu 285
DB 985 GTTTCTCTTAGTGAATCCATAGTAGTTATTCATTTGGATGGAGCAGCTCCCTCACTTTGAA 1044
QY 286 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPro 305
DB 1045 ACCTTCGAATAGCCCGAGGAGCTGCCTTTTCATATTTTCCAGGTTATTGATAAGAAACCC 1104

QY 306 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 325
DB 1105 AGTATAGGTAACTTTTCCACAGCTGGATATAAACCTGAAATCCATAGAGAACTGTGGAA 1164
QY 326 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu 345
DB 1165 TTTAAAAATGTTTCTTTCAATTATCCATCAAGGCCATCTATCAAGATTCTGAAAGGTCTG 1224
QY 346 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 365
DB 1225 AATCTCGGAATTAAGTCTGGAGAGACAGTCGCTTGGTGGTCTCAATGGCAGTGGGAAG 1284
QY 366 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 385
DB 1285 AGTACGGTAGTCAGGCTTCTCGAGAGTTATATGATCCGGATGATGGCTTATCATGGTG 1344
QY 386 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal 405
DB 1345 GATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATTATGGAGTGGTT 1404
QY 406 SerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAsp 425
DB 1405 AGTCAAGAGCCTGTTTGTTCGGGACCACCATCAGTAAACAATATCAAGTATGGACGAGAT 1464
QY 426 AspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIle 445
DB 1465 GATGTGACTGATGAAGAGATGGAGAGCAGCAGCAGGAGGAGCAAAATCGGTATGATTTTATC 1524
QY 446 MetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGly 465
DB 1525 ATGGAGCTTTCCTAAATAATTTAATACATTTGTTAGGGGAAAAAGAGAGCTCAATAGTGGGA 1584
QY 466 GlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIle 485
DB 1585 GGGCAGAAAACAGAGGATCGCAATTCGTGCTGCTTGTAGTTCGAAACCCCAAGATTCGATT 1644
QY 486 LeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeu 505
DB 1645 TTAGATGAGCTACGTCTGCTCCCTGGATTCAGAAAGCAAGTCAGCTGTTCAAGTGCACCTG 1704
QY 506 GluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIleArg 525
DB 1705 GAGAAGCGCAGCAAAAGTTCGACTCAATCTGCTGGTAGCACACCGAGCTTCTACTACTTCCA 1764
QY 526 SerAlaAspIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAla 545
DB 1765 AGTGCAGATTGATTGTGACCTTAAAGGATGGAATGCTGCGGAGAAAGGAGCGCATGCT 1824
QY 546 GluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIleLysLys 565
DB 1825 GAATTAATGGCAAAACAGAGCTTATATTATTCCTTGTGATGTCACAGGATATTAATAAAA 1884
QY 566 AlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuPro 585
DB 1885 GCTGATGAACAGATGAGTCAATGACATATCTACTGAAAGAAAGCAACCACTCACTTCCT 1944
QY 586 LeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThrGln 605
DB 1945 CTGCACTCTGTGAAGAGCATCAAGTCAGACTTCATTGACAAGGCTGAGGAATCCACCCAA 2004
QY 606 SerLysGluIleSerLeuProGluValSerLeuLysIleLeuLysLeuAsnLysPro 625
DB 2005 TCTAAAGAGATAGTCTTCTCTGAAGTCTCTCTATTAATAAATTTTAAAGTTAAACAAGCCT 2064
QY 626 GluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPro 645
DB 2065 GAATGSCCTTTTGTGCTCTCTGGGACATTTGGCTTCTGTTCTAAATGGAACTGTTTCATCCA 2124
QY 646 ValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThr 665
DB 2125 GTATTTTCCATCATCTTTGCAAAATTTATAACCATGTTTGGAAATTAATGATAAAACACA 2184

QY 666 LeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPhe 685
DB 2185 TTAAGCATGATGATGCGAATAATTTATTCATGATATTCGTTGGGTGTTATTTGCTTT 2244
QY 686 ValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArg 705
DB 2245 GTCAGTTATTTTCATG----- 2259
QY 706 LeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLys 725
DB 2260 -----CAGGATATGCTCTGGTTTGGATGAAAG 2286
QY 726 GluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGly 745
DB 2287 GAAACAGACAGGAGGCTTGACAAATATTAGCCATAGATATAGACAAAATTCAGGA 2346
QY 746 AlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerVal 765
DB 2347 GCACAGGTTCCAGGATTGGCTCTTAACACAAATGCAACTACATGGACTTTTCAGTT 2406
QY 766 IleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProVal 785
DB 2407 ATCATTTCTTTATATATGATGGAGATGACATTCCTGATCTGAGTATTGCTCCAGTA 2466
QY 786 LeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLys 805
DB 2467 CTTGCCGTGACAGGATGATTGAACCGCAGCAATGACTGGATTTCGCAACCAAGATAAG 2526
QY 806 GlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuIleAsnIleArgThrIle 825
DB 2527 CAAGAACTTAAGCATCTGGAAGATAGCAACTGAAGCTTTGGAGATATATACGACTATA 2586
QY 826 ValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGln 845
DB 2587 GTGCTATTAAACAGGAAAGGCTTCGAGCAATGTATGAAGATGCTTCAGACTCAA 2646
QY 846 HisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAla 865
DB 2647 CACAGAAATACCTCGAAGAAAGCACAGATATTATGGAAGCTGTTATGCATTCAGCCATGCC 2706
QY 866 PheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyValaTyrLeuIleGlnAlaGly 885
DB 2707 TTTATATATTTTGCCTATGCGGAGGCTTCGATTTGGAGCCATTATTAATCAAGCTGGA 2766
QY 886 ArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaIleAlaTyrGlyAlaMetAla 905
DB 2767 CGATACCCAGGAGGCGATGTTTCATAGTTTTTACTGCAATTCATATGAGAGCTATGGCC 2826
QY 906 IleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHis 925
DB 2827 ATCGGAGAAACGCTGCTTTTGGCTCTCTGAATATTCAAAGCCAAATCGGGGGCTGGCAT 2886
QY 926 LeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluLysLys 945
DB 2887 CTGTTTGGCTTGTGGAAGAAACCAATATATAGACGCGCGAGTCAAGAAGGGAAGAAAG 2946
QY 946 ProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArg 965
DB 2947 CCAGACACATGTGAGGGAATTTAGAGTTTCGAGAGTCTCTTTCTTATCCATGTCCG 3006
QY 966 ProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAla 985
DB 3007 CCAGATGTTTTCATCTCCGTGGCTTATCCCTCAGTATTGAGCGAGGAAAGACAGTAGCA 3066
QY 986 PheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuGlnArgLeuTyr 1005
DB 3067 TTTGTGGGAGCAGCGCTGTGGGAAGACACTTCTGTTCAACTTCTGCAGAGACTTTAT 3126
QY 1006 AspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGln 1025
DB 3127 GACCCCGTCAAGGACAAAGTCTGTTTGTGGTGTGATGCAAAAGAAATTTGAATGTACAG 3186
QY 1026 TrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIle 1045

DB 3187 TGGCTCCGTTCCCAATAGCAATCGTTCTCAAGAGCCCTGTGCTCTTCACTGCACAT 3246
QY 1046 AlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGlu 1065
DB 3247 GCTGAGAACATCGCTATGGTGACAAACAGCCGTGTGGTCCATTAGATGAGATCAAGAA 3306
QY 1066 AlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThr 1085
DB 3307 GCGCAAAATGCAAGCAATATCCATTTCTTTTATTTGAAGGTCTCCCTGAGAAATACAACACA 3366
QY 1086 GlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAla 1105
DB 3367 CAAGTTGGACTGAAAGGACAGACTTTCTGCGGCCAGAAAACAAAGACTAGCTATTGCA 3426
QY 1106 ArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAsp 1125
DB 3427 AGGGCTCTTCTCAAAACCCAAATTTTATTTGGATGAGGCCACTTCAGCCCTCGAT 3486
QY 1126 AsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCys 1145
DB 3487 AATGACAGTGAGAAGTGGTTGAGCATGCCCTTGATAAAAGCCAGGACGGGAAGGACATGC 3546
QY 1146 LeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHis 1165
DB 3547 CTAGTGGTCACACAGGCTCTCTGCAATTCAGAACGCGAGATTTGATAGTGGTCTGCAC 3606
QY 1166 AsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyr 1185
DB 3607 ATGGAAAGTAAAGAACACAGGACTCATCAAGACTCTCTGAGAAATCGAGACATATAT 3666
QY 1186 PheLysLeuValAsnAlaGlnSerValGln 1195
DB 3667 TTTAAGTTAGTGAATGCACAGTCAGTGCAG 3696
RESULT 4
AX339030 AX339030 3177 bp DNA linear PAT 09-JAN-2002
LOCUS Sequence 12 from Patent WO0194400.
DEFINITION AX339030
ACCESSION AX339030
VERSION AX339030.1 GI:18129122
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Frank, M.H. and Sayegh, M.H.
TITLE A gene encoding a multidrug resistance human p-glycoprotein
homologue on chromosome 7p15-21 and uses thereof
JOURNAL Patent: WO 0194400-A 12 13-DEC-2001;
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES
source Location/Qualifiers
1..3177
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 970 a 601 c 735 g 870 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 3177
Score: 5159.50 Matches: 1031
Percent Similarity: 97.45% Conservative: 0
Best Local Similarity: 97.45% Mismatches: 0
Query Match: 85.82% Indels: 27
DB: 6 Gaps: 1
US-09-873-409-6 (1-1195) x AX339030 (1-3177)
QY 165 MetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAla 184
DB 1 ATGGTCATCTCATTCAGGAAATTAAGTGCCTATTCCAAAGCTGGGCTGGCA 60


```
Db 2221 GGAGCCTATTAAATCAAGCTGACGAATGACCCCGAGAGGGCAGTTCATAGTTTTTACT 2280
Qy 898 AlaileAlaTyrglyAlaMetAlaileGlyLysThrLeuValLeuAlaProGluTyrsr 917
Db 2281 GCAATTTGCATATGAGCCTATGGCCATCGGAAACCGCTCGTTTGGCTCTCTGAATATCC 2340
Qy 918 LysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAsp 937
Db 2341 AAGCCAAATCGGGGCTCGCATCTGTTTGGCTTGTGGAAAGAAACCAATATAGAC 2400
Qy 938 SerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGlu 957
Db 2401 AGCCGCGTCAGACGGAAGAAAGCCAGACACATGTGAGGGAATTTAGAGTTTCAGAA 2460
Qy 958 ValSerPhePheTyrglyProCysArgProAspValPheileLeuArgGlyLeuSerLeuSer 977
Db 2461 GTCTCTTTCTTCTATCCATGTCGCCCAGATGTTTTCATCTCGTGGCTTATCCCTCAGT 2520
Qy 978 IleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSer 997
Db 2521 ATTGAGCGAGGAAGACAGTAGCATTTGTGGGAGCAGCGGCTGTGGGAAAGCACTTCT 2580
Qy 998 ValGlnLeuLeuGlnArgLeuTyrglyProValGlnGlyGlnValLeuPheAspGlyVal 1017
Db 2581 GTTCAACTTCTGAGAGACTTTATGACCCCGTCGACAGGCAAGTGTCTGTGATGGTG 2640
Qy 1018 AspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGlu 1037
Db 2641 GATGCAAGAGATTGATGTACAGTGGCTCCGTTCCCAATAGCAATCGTTCTCAGAG 2700
Qy 1038 ProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrglyAspAsnSerArgVal 1057
Db 2701 CCTGTCTCTTCAACTGCAGCATTTGTGAAACATCGCCTATGCTGACACAGCCGTGTG 2760
Qy 1058 ValProLeuAspGluLeuLysGluAlaAlaAsnAlaAsnIleHisSerPheIleGlu 1077
Db 2761 GTGCCAATTAGATGAGATCAAGAGCGCAATGCGCAAAATATCCATCTCTTTATTGAA 2820
Qy 1078 GlyLeuProGluLysTyrglyAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGly 1097
Db 2821 GGTTCTCTCTGAGAAATACACACACAGTTGGACTGAAAGGACACAGCTTCTGCGGCG 2880
Qy 1098 GlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeu 1117
Db 2881 CAGAAACAAAGACTAGCTATTGCAAGGGCTCTCTCCAAAACCCCAAAATTTTATTGTTG 2940
Qy 1118 AspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAsp 1137
Db 2941 GATGAGGCCACTTCAGCCCTCGAATATGACAGTGAAGAGTGGTTGAGCATGCCCTTGAT 3000
Qy 1138 LysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsn 1157
Db 3001 AAGCCAGGAGGGAAGGACATCGCTAGTGGTCATCTCAGAGGCTCTCTGCANTTCAGAAC 3060
Qy 1158 AlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlu 1177
Db 3061 GCAGATTTGATAGTGGTTCTGCAACATGGAAGAGTAAGAAAGCAAGAACTCATCAAGAG 3120
Qy 1178 LeuLeuArgAsnArgAspIleTyrglyPheLeuLeuValAsnAlaGlnSerValGln 1195
Db 3121 CTCTCTGAGAAATCGAGACATATATTTTAAAGTTAGTGAATGAGTGCAGTGCAGTGCAG 3174

RESULT 5
AX339028
LOCUS AX339028 2856 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 10 from Patent WO0194400.
ACCESSION AX339028
VERSION AX339028.1 GI:18129120
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

```
REFERENCE 1
AUTHORS Frank, M.H. and Sayegh, M.H.
TITLE A gene encoding a multidrug resistance human p-glycoprotein
JOURNAL homologue on chromosome 7p15-21 and uses thereof
PATENT: WO 0194400-A 10 13-DEC-2001;
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES
source location/Qualifiers
1..2856
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 888 a 540 c 652 g 776 t
ORIGIN
Alignment Scores:
Pred. No.: 8.1e-288 Length: 2856
Score: 4561.50 Matches: 912
Percent Similarity: 97.96% Conservative: 0
Best Local Similarity: 97.96% Mismatches: 1
Query Match: 75.87% Indels: 18
DB: 6 Gaps: 1
US-09-873-409-6 (1-1195) x AX339028 (1-2856)
Qy 266 ValPhePheSerValIleHisSerSerTyrglyCysIleGlyAlaAlaValProHis-PheGI 285
Db 114 GTTTTCTTTAGTGAATCCATAGCAGTTATTGTCATTGGAGCAGCAGTCCCTCATTTATT-- 171
Qy 285 uThrPheAlaIleAlaAraGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPr 305
Db 172 -----GATNAGAAACC 182
Qy 305 oSerIleAspAsnPheSerThrAlaGlyTyrglyProGluSerIleGlyGlyThrValGI 325
Db 183 CAGTATAGATAAATCTTTCCACAGCTGATATAAACTGAATCCATAGAGGAACCTGGA 242
Qy 325 uPheLysAsnValSerPheAsnTyrglyProSerArgProSerIleLysIleLysGlyLe 345
Db 243 ATTTAAAAATGTTCTTTCAATTTATCCATCAAGACCATCTATCAAGATTTCTGAAGGTCT 302
Qy 345 uAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLy 365
Db 303 GAATCTCAGAAATTAAGTCTGGAGACAGTCGCTTGGTCTCANTGGCAGTGGAA 362
Qy 365 sSerThrValValGlnLeuLeuGlnArgLeuTyrglyProAspAspGlyPheIleMetVa 385
Db 363 GAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCGATGATGGCTTTATCATGGT 422
Qy 385 lAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrglyArgAspHisIleGlyValVa 405
Db 423 GGATGAGAAATGACATCAGAGCTTTAAATGTCGGCATTTATCGAGACCATATTGGAGTGGT 482
Qy 405 lSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrglyArgAs 425
Db 483 TAGTCAAGAGCCCTGTTTGTTCGGGACCCACCATCAGTAAACAATATCAAGTATGGACAGA 542
Qy 425 pAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrglyPheIle 445
Db 543 TGATGTGACTGATGAAGAGATGGAGAGACAGCAAGGAAAGCAAAATCGCTATCATTTAT 602
Qy 445 eMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGI 465
Db 603 CATGGAGTTTCTTAATAAATTTAATACATTGGTAGGGAAAAAGAGAGCTCAAAATGAGTGG 662
Qy 465 yGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuII 485
Db 663 AGGGCAGAAACAGAGATCGCAATTCCTCGTCTCTAGTTTCGAAACCCCAAGATTCGTAT 722
Qy 485 eLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaIle 505
Db 723 TTTAGATGAGGCTACGTCCTCCCTGGAATTCAGAAACAGTCAAGTGTTCAGCTGCACCT 782
Qy 505 uGluLysAlaSerLysGlyArgThrThrIleValAlaHisArgLeuSerThrIleAr 525
|||||
```


Db 783 GGAGAGGCGAGCAAGGTCGGACTCAATCGTGGTAGCACACCGACTTCTTACTATTTCG 842
QY 525 gSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAl 545
Db 843 AAGTGCAGATTGATGTGACCCCTAAAGGATGGAATGCTGGCGGAGAAAGGACACATGC 902
QY 545 aGluLeuMetAlaLysArgGlyLeuTyfyrSerLeuValMetSerGlnAspIleLysLy 565
Db 903 TGAACATAATGGCAAAACGAGGTCTATATTATTCATTGTGATGCACAGGATATAAAA 962
QY 565 sAlaAspGluMetGluSerMetThrTyfyrSerThrGluArgLysThrAsnSerLeuPr 585
Db 963 AGCTGATGAACAGATGAGTCAATGACATATCTTACTGAAAGAAAGACCAACTCATCTTC 1022
QY 585 oLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluLysSerThrGl 605
Db 1023 TCTGCACCTCTGTGAGGAGCATCAAGTCAGACTTCATTGACAAAGGCTGAGGAATCCACCCA 1082
QY 605 nSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysPr 625
Db 1083 ATCTAAAGAGATAAGTCTTCTCTGAAGTCTCTCTATTAAAAAATTTTAAAGTTAAACAAGCC 1142
QY 625 oGluTrpPropheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPr 645
Db 1143 TGAATGGCCCTTTTGTGGTCTCGGGACATTTGGCTTCTGTTCTAAATGGAACTGTTTCATCC 1202
QY 645 oValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrTh 665
Db 1203 AGTATTTTCCATCATCTTTGCAAAAATATTAACCAATGTTTGGAAATATGATTAACCCAC 1262
QY 665 rLeuLysHisAspAlaGluIleTyfyrSerMetIlePheValIleLeuGlyValIleCysPh 685
Db 1263 ATTAAGCATGATGCAGAAATTTATTCATGATATTCGTTCATTTGGGTGTTATTGGCTT 1322
QY 685 eValSerTyfyrPheMetGlnGlyLeuPheTyfyrGlyArgAlaGlyGluIleLeuThrMetAr 705
Db 1323 TGTCAAGTATTTCATGCGAGGATTTATTTACGCGAGCAGCAGGGGAAATTTTAAACGATGAG 1382
QY 705 gLeuArgHisLeuAlaPheLysAlaMetLeuTyfyrGlnAspIleAlaTrpPheAspGluLy 725
Db 1383 ATTAAGACACTTGGCCTTCAAGCAATGTTATATCAGGATATTGCCTGGTGTGATGAAA 1442
QY 725 sGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGl 745
Db 1443 GGAAAAACAGCACAGGAGCCTTGACAAATATTAGCCATAGATATACCAAAATTCAGG 1502
QY 745 yAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerVa 765
Db 1503 AGCAACAGGTTCCAGGATTTGGCGTCTTAAACACAAAATGCAACTAACATGGGACTTTTCAGT 1562
QY 765 lIleIleSerPheIleTyfyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProVa 785
Db 1563 TATCATTTCTTTATATATGATGGAGATGACATTTCTGATTTCTGATGATTTGCTCCAGT 1622
QY 785 lLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLy 805
Db 1623 ACTTGGCGTGCAGGAATGATTTGAAACCGCAGCAATGACTGGATTTGCCAACAAAGATAA 1682
QY 805 sGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrII 825
Db 1683 GCAAGAACTTTAAGCATCTCGAAGATAGCAACTGAAAGCTTTGAGAAATATACGTACTAT 1742
QY 825 eValSerLeuThrArgGluLysAlaPheGluGlnMetTyfyrGluGluMetLeuGlnThrGl 845
Db 1743 AGTGTCTATTAAACAGGGAAGAAAGCCCTTCGAGCAAAATGATGAAGAGATGCTTCAGACTCA 1802
QY 845 nHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyfyrAlaPheSerHisAl 865
Db 1803 ACACAGAAATACCTCGAAGAAAGCACAGATATTATTGGAAGCTGTTATGCATTCAGCCATGC 1862
QY 865 aPheIleTyfyrPheAlaTyfyrAlaAlaGlyPheArgPheGlyAlaTyfyrLeuIleGlnAlaGl 885
Db 1863 CTTTATATATTTTGCCTATGATGAGCAGGGTTTCGATTTTGGAGCCCTATTATTAATTCAAGCTGG 1922

QY 885 yArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyfyrGlyAlaMetAl 905
Db 1923 ACGAATGACCCAGAGGGCATGTTTCATAGCTTTTCTCAATTTGCATATGGAGCTATGGC 1982
QY 905 alleGlyLysThrLeuValLeuAlaProGluTyfyrSerLysAlaLysSerGlyAlaAlaHi 925
Db 1983 CATCGAAAAACGCTCGTTTTGGCTCTGAAATATTCCAAGCCAAATCGGGGCTCGCA 2042
QY 925 sLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluLysLy 945
Db 2043 TCTGTTTGGCTTGTGGAAAAAGAACCAATATAGACACCGCAGTCACAGAGGAAAAA 2102
QY 945 sProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyfyrProCysAr 965
Db 2103 GCCAGACACATGTGAAGGAAATTTAGAGTTTCAGAAAGTCTCTTTCTTATCATGTGCG 2162
QY 965 gProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAl 985
Db 2163 CCAGATGTTTTCATCCTCCGTGGCTTATCCCTCAGTATTGAGCGAGGAAAGACAGTAGC 2222
QY 985 aPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTy 1005
Db 2223 ATTTTGGGAGCAGCGCTGTGGAAAAAGCACTTCTGTTCAACTTCGACAGACTTTA 2282
QY 1005 rAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGl 1025
Db 2283 TGACCCCGTGCAAGGACAAAGTGTGTTTCATGTTGATGCAAAAAGAAATGAATGTACA 2342
QY 1025 nTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIl 1045
Db 2343 GTGGCTCCGTTCCTCAAAATAGCAATCGTTCTCAAGAGCCCTGTCTCTTCAACTGCA 2402
QY 1045 eAlaGluAsnIleAlaTyfyrGlyAspAsnSerArgValValProLeuAspGluIleLysGl 1065
Db 2403 TGTGTAGAACATCGCCTATGGTGACAAACAGCCGTGGTGCCATTAGATGAGATCAAGA 2462
QY 1065 uAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyfyrAsnTh 1085
Db 2463 AGCCGCAATGCAGCAAAATATCCATTTTATTGAAGGTCTCCCTGAGAAATACAACAC 2522
QY 1085 rGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAl 1105
Db 2523 ACAAGTTGACTGAAAGGAGCAGCTTTCTGGCGCCAGAAAACAAAGACTAGCTATTGC 2582
QY 1105 aArgAlaLeuLeuGlnLysProLysIleLeuLeuAspGluAlaThrSerAlaLeuAs 1125
Db 2583 AAGGCTCTTCTCCAAAAACCCAAATTTTATTGTTGGATGAGCCACTTCAGCCCTCGA 2642
QY 1125 pAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCy 1145
Db 2643 TAATGACAGTGAAGAGTGGTTTCAGCATGCCCTTGTATAAAGCCAGGACGAGGACATG 2702
QY 1145 sLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHi 1165
Db 2703 CCTAGTGGTCACTACAGGCTCTCTGCAATTCAGAACGCGAGATTTGATAGTGTGCTGCA 2762
QY 1165 sAsnGlyLysIleLeLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTy 1185
Db 2763 CAATGAAAGATTAAGAAACAAAGAACTCATCAAGAGCTCTCTGAGAAATCGACATATA 2822
QY 1185 rPheLysLeuValAsnAlaGlnSerValGln 1195
Db 2823 TTTTAAAGTAGTGAATGCACAGTCAGTCAGTCAG 2853
RESULT 6
CRUPGP1
LOCUS C.griseus p-glycoprotein (isoform class I) mRNA linear ROD 22-JUN-1995
DEFINITION M60040
ACCESSION M60040.1 GI:191164
VERSION multidrug resistance; p-glycoprotein; transmembrane protein.
KEYWORDS C.griseus adult liver and ovary, cdna to mRNA.

[illegible]


```

Db 3709 CTGATACAGAGAGTGAAGAGTGTCCAGAGCTCTGGACAAAGCCAGAGAGCCGC 3768
Qy 1144 ThrCysLeuValThrHisArgLeuSerAlaIleGlnAsnAlaSerLeuLeuValVal 1163
Db 3769 ACTGATGTGATCGTCAACCGCTGTCCACCATCCAGAACCGACACTGTGCTGGTG 3828
Qy 1164 LeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsp 1183
Db 3829 ATTCAGATGGCAAGTCAAGGAGCAGTGGCACCCAGCAGCAGCTGTGTCGACAGAAAGC 3888
Qy 1184 IleTyrPheLysLeuValAsnAlaGlnSer 1193
Db 3889 ATCTATTCTCCATGTGTCAGTGTGCAGGCT 3918

RESULT 7
CRUPGP1165 CRUPGP1165 4304 bp mRNA linear ROD 27-APR-1993
LOCUS Chinese hamster p-glycoprotein mRNA (clone ADX165), complete cds.
ACCESSION M59253
VERSION M59253.1 GI:191154
KEYWORDS multidrug resistance glycoprotein; p-glycoprotein.
SOURCE Chinese hamster, cDNA to mRNA, clone ADX165.
ORGANISM Cricetulus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
Cricetulus.
REFERENCE 1 (bases 1 to 4304)
AUTHORS Devine,S.E., Husain,A., Davide,J.P. and Melera,P.W.
TITLE Full length and alternatively spliced pgpi transcripts in
multidrug-resistant Chinese hamster lung cells
JOURNAL J. Biol. Chem. 266 (7), 4545-4555 (1991)
MEDLINE 91154265
PUBMED 1671863
FEATURES
     source
         1..4304
         /organism="Cricetulus sp."
         /db_xref="taxon:10031"
         /cell_type="fibroblast"
         /tissue_type="lung"
     gene
         1..4304
         /gene="pgp-1"
         /codon_start=1
         /product="p-glycoprotein"
         /protein_id="AA37004.1"
         /db_xref="GI:191155"
         /translation="MEFEDEFGRKDKNFKMKRKSKEKKEKPPVSVFTFPRVAGM
         LDRMLVGTALAIHGVALPLMLVFGDMTDSFASVGNIPTNATNNATQVNASDIFG
         KLEEMTVAYTYTGTGAGLVAVIYGVSWCLAAAGROIHKROKFFHAIHQEIGWF
         DVHVGELNRLTDVSKINEGDKIMGFQMATFFGGFIIGTRGWKLTVLILAI
         SPVLGSLGIWAKILSSFTDKELQAYKAGAVAEVLAIKRVIAFGGKKELRYNN
         NLEEARLKIANTANISMGAFLLIYASALAFWYGLSVLSKEYSIGQVLTFFFA
         VLIAPFSIQGASPNIEAFANARGAAEYFNIDNPSIDFSKNGYKPDNIKNLEFK
         NLFHSPSRDKQVILKGLNKVSGOTVALVNGSCGKSTTVLLQLRLVDPTGVVSI
         DQDQRTINVRILREITIGVSVSEPLVFTATTIENIRYGRNVMTDEIEKAVEANAYD
         FIMKLPKFTDLVGRGQLSGQKORIALALVRLNPKILLIDENATSLDITSENAVY
         QALDKAREGRTIIVIAHRLSTVIRADIIAGFDGVIYVEQGNHLEMKGIYFKLVM
         TQTAGNELEVEGESKNEIDNLMSKDSASSLIRRRSTRSIRGPHDQRLKSTK
         EALDEVPPISEFIRLKNLSNPPYVVGIFCAIVNGALQAPAFSIFSKVGVFTRNT
         DDETRKDNLSLLPLIGVLSFITFFLQGTFFGKAGELTKRLRYVFKSMLRQDV
         SWFDNPKNTGALTFLANDAGVKGATGARLVATONLGTGIIISLIYGMQTLT
         LLALVPIIAGVEMKMSQALKDKKEGSKKIATEIENRTVSVLSRQKFE
         NMYAQSLPYRNALKHAKVGFITFTQAMMIFSTAACFRFGALVARELMEFNVL
         LVFSIVFGAMQVQSSPAPYAKAKVSASHIIMIEKVPISIDSYSGLKPNTLEG
         NVKFNVEVYTRPDIPVLQGLNLEVKKGOTLALVSSCGSKSTVQLLERFYDFMA
         GTVFLDGKEVNOLNQMWRHLGIVSOEPLIFDCSTAEINAYGDNRSRVVSOEIERAA
         KEANTHOFESLPDKYNTRVGDKGTQLSGQKORIALALVROPHILLLDRATSLD
         TESEKVVQALDKAREGRTIIVIAHRLSTIQLNADLLIVLVQNGKVKHGHQQLAQKG
         IYFSWVQAGAR"
     map
         1275 a 895 c 1031 g 1103 t
         Map position 1q26.

```

BASE COUNT
ORIGIN

Alignment Scores:
Pred. No.: 3,04e-223 Length: 4304
Score: 3573.00 Matches: 679
Percent Similarity: 75.37% Conservative: 248
Best Local Similarity: 55.20% Mismatches: 259
Query Match: 59.43% Indels: 44
DB: 10 Gaps: 8

US-09-873-409-6 (1-1195) x CRUPGP1165 (1-4304)

Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 258 ATGCTGGTGGGACTCTCGCAGCTATTATCATGGAGTTGCATCCCGCTTATGATGCTG 317
Qy 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsn----- 38
Db 318 GTGTTGGAGACATGACACAGCTTTGCAAGTGTAGGAACAATCCCAACAATGCTACT 377
Qy 38 ----- 38
Db 378 AATAATGCCACTCAAGTCAATGCTTCAGACATCTTTGGCAAACTGGAGGAAGAAATGACC 437
Qy 39 ThrTyrSerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIle 58
Db 438 ACGTACGCCTAT-----TACTACACAGGATTTGGTCTGCTGCTGCTTATA 482
Qy 59 PheGlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIle 78
Db 483 GTTGCTTACATTCAGTTTCATTTTGGTGCCTGGCAGCTGGGAAGACAATACACAAAATT 542
Qy 79 ArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAsp 98
Db 543 AGACAGAAGTTTTTTCATGCTATAATGAATCAGGAGATAGGCTGGTTGACGTGCAGCAC 602
Qy 99 IleGlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGly 117
Db 603 GTTGGGGAGCTCAACACCGCGCTCACAGATGATGTCTCCAAAATTAAACGAAGAAATTGGT 662
Qy 118 AspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGly 137
Db 663 GACAAAATTGGATGTTCTTTTTCAGGCAATGGCAACAATTTTGGTGGTTTATATAGGA 722
Qy 138 LeuValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAla 157
Db 723 TTTTACTCGAGGCTGGAAGCTAACCTTTGTGATTTTGGCCATCAGCCCTGTTCTTGACATG 782
Qy 158 SerAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyr 177
Db 783 TCAGCTGGTATTTTGGCAAGATATTATCTTCTATTACTGATAAAGAACTTCAGGCATAT 842
Qy 178 SerIleAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPhe 197
Db 843 GCAAAAGCTGGAGCAGTGTCTGAAGAAGTCTTAGCAGCCATCAGAACTGATGATTCATTT 902
Qy 198 ArgAlaGlnGlyLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPhe 217
Db 903 GAGGACAAAGAAAGAACTTGAAGGTACACAACTTTAGAAAGAGCCAAAGACTT 962
Qy 218 GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsn 237
Db 963 GGGATAAAGAAAGCGATCACAGCAACATTTCCATGGGCGCAGCTTTCCTTCTTATCTAT 1022
Qy 238 GlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerIleLeuAsnGlyGluProGly 257
Db 1023 GCATCATATGCTCGGCCCTTGGTATGAACGCTCTTGGTTCATCTCAAAAGAA----- 1076
Qy 258 TyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIle 277
Db 1077 TATCTATTGGACAAGTGTCTACTGTCTTCTTGTGCTATTATATTCGCCCATTCAGTATT 1136
Qy 278 GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 297

Db	1137	GGACAGCGCATCTCCAATAATCGAAGCCTTTGGCAATGCAAGAGGAGCAGCCTTATGAATC	1119
Qy	298	PheGlnValIleAspLysProSerIleAspAsnPheSerThrAlaGlyTyrLysPro	317
Db	1197	TTCAACATAATTGATAAAGCCAGTATTGACAGCTTCTCAAGAATGGGTACAAACCA	1256
Qy	318	GluserIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgPro	337
Db	1257	GACAACATTAAAGCAAAATTGGAAATTCAAAAATATTCACTTCAGTTACCCATCTCGAAAA	1316
Qy	338	SerIleLysIleLeuLysGlyLeuAsnLeuAArgIleLysSerGlyGluThrValAlaLeu	357
Db	1317	GAGCTTCAGATCTTTGAAGGCGCTCAACCTGAAGGTGAGAGCGGACAGACAGTGGCCCTG	1376
Qy	358	ValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAsp	377
Db	1377	GTGGCAACAGTGGCTGTGGGAAGAAGCACCATCTGTCCAGCTGCTGCAGAGGCTCTACGAC	1436
Qy	378	ProAspAspGlyPheIleMetValAspGluAsnAspIleAArgAlaLeuAsnValArgHis	397
Db	1437	CCACAGAGGCGGTGTGTCAGTATCGACGCGCAGGACATCAGGACCATCAATGTGAGGTAT	1496
Qy	398	TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSer	417
Db	1497	CTCGGGAATCATTTGGGTGCTGAGTCAGGAACCTGTGTGTTTGGCCACCACATTCGCT	1556
Qy	418	AsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArg	437
Db	1557	GAAAACATTGCTATGGCCGAGAAAATGTCACTATGGAGGAGATCGAGAAAGCTGTCAAG	1616
Qy	438	GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly	457
Db	1617	GAGCCCAATGCCATTGACTTCATCATGATAAGCTGCCCATATAAATTGACACTCTTGTGGT	1676
Qy	458	GluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu	477
Db	1677	GAGAGAGGACACAGCTGAGTGGAGACAGAAACAGAGAAATCGCCATTGCTCTGTCGCCCTG	1736
Qy	478	ValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer	497
Db	1737	GTCCGCAACCCCAAGATCTCTTTTGGATGAGGCGAATCAGCCTTGGACACAGAAGT	1796
Qy	498	LysSerAlaValGlnAlaLeuGluLysAlaSerLysGlyArgThrThrIleValVal	517
Db	1797	GAGCCCGTGTTCAGGCTGCTCGGATAAGGCTAGAGAAGCGCGGACTACCATTTGTGATA	1856
Qy	518	AlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMet	537
Db	1857	GCCCAACCGCTGTCTACAGTTTCGAAATGCTGCATAGCTGGGTTTGGTGGTGGTGC	1916
Qy	538	LeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeu	557
Db	1917	ATTGTGAGCAGAGAAATCATGAAGAGCTCATGAGAGAGAGGCGGATTTACTTCAAACTT	1976
Qy	558	ValMetSerGln-----AspIleLysLysAlaAsp	567
Db	1977	GTCAATGACACAGACAGCGAGAAATGAAATTGAATTAGGAATGAAGTTGGTGAGCTTAAA	2036
Qy	568	GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeu-----Pro	585
Db	2037	AATGAAATTGATAAATTAGACATGTCTTCAAAAGATTTCAGCATCCAGTCTAATCAGAAGG	2096
Qy	586	LeuHisSerValLysSerIleLysSerAppPheIleAspLysAlaGluLysSerThr---	604
Db	2097	AGATCAACTCCGAGAAGTATCCGTGGACCCACATGACCAAGACAGGAAGCTTAGTACCAA	2156
Qy	605	GlnSerLysGlnIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLys	624
Db	2157	GAGCGCTTGAATGAGATGTACCTCCAATTTCTTTTGGCGTATCTTAAGTGAATCA	2216
Qy	625	ProGluTrpPropheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHis	644
Db	2217	TCTGAAATGGCCATATTTGTTGGTTGGTATATTCTGTGCCATAGTAATGAGACCTTGCAA	2276

Qy	645	ProValPheSerIlellePheAlaLysIlelleThrMetPhe---GlyAsnAsnAspLys	663
Db	2277	CCAGCATCTCTCAATAATATTTCTAAGGTGTAGGGGTTTTTCACAAGAATAATACATGATGAT	2336
Qy	664	ThrThrLeuLysHisAspAlaGluIleThrSerMetIlePheValIleLeuGlyValIle	683
Db	2337	GAACACCAACGACATGATAGCAACTTGTTTTCTTATTTGTTTCTGATCTCTTGGGGTCATT	2396
Qy	684	CysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThr	703
Db	2397	TCCTTTTATTACATTTTCTCTCAGGCGCTTCACATTTGGCAAGCTGGAGAGATCCTCACC	2456
Qy	704	MetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaThrPheAsp	723
Db	2457	AAGCGATCTCGATACATCGTTTTCAAATCCATGCTGAGACAGATGATGAGCTGGTTTGAT	2516
Qy	724	GluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIle	743
Db	2517	AACCTTAATAAACACCACTGGAGCATTTGACCAACGAGCTCGCCACGATGCTGGTCAAGTT	2576
Qy	744	GlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeu	763
Db	2577	AAAGGGGCTACAGAGCCGACACTTGCTGTCATTATCCCAAGAAATATAGCAAAATCTTGGGACA	2636
Qy	764	SerValIleIleSerPheIleTyrGlyTyrTpGluMetThrPheLeuIleLeuSerIleAla	783
Db	2637	GGATCATCATATCCCTTAATCATGCTGGCAGTTGCACATCTTACTCTTAGCAATTGTC	2696
Qy	784	ProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLys	803
Db	2697	CCCATATTCGAATAGCAGGAGTGGTGTAGATGCAAAATGTTCTCTGGACAAGACACTAAAA	2756
Qy	804	AspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArg	823
Db	2757	GATAAGAAGGAGCTAGAAGGTTCTGGGAAGATGTCTACTGAAGCAATAGAGAACTTCGCG	2816
Qy	824	ThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGln	843
Db	2817	ACTGCTGCTCTTTGACTCGGGAGCAGAAGTTTGAATAATATGATGCCAGAGCTTGCGAG	2876
Qy	844	ThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSer	863
Db	2877	ATACCATACAGAAATGCTCTGAAGAAGCACACGCTTTGGGATCACATCTCTCTTACC	2936
Qy	864	HisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGln	883
Db	2937	CAGGCCATGATGATTTTTCGTATGCTGCTTGTTCGGTTTGGTGTCTTACTTGTGTGCA	2996
Qy	884	AlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAla	903
Db	2997	CGGGAATTAATGACATTTGAAATGTCTATTAGTATTCCTCAGCTATTGTCTCTTGTGTGCC	3056
Qy	904	MetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAla	923
Db	3057	ATGCGAGTTGGCAGGTGACTTCATTTGCTCTGACTATGCCAAGACCAAGAGTGTGCGCA	3116
Qy	924	AlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGly	943
Db	3117	TCGCATATCATCATGATCATTTGAAAGAAGTCCCTTCCATTGACAGCTACAGCACCGGAGGC	3176
Qy	944	LysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrPro	963
Db	3177	CTGAAGCGCTAATCATTTGGAAGGAATGTGAAATTTAATGAAGTCGTGTCAACTTACCCC	3236
Qy	964	CysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThr	983
Db	3237	ACCGGGCTGNATCCACCGTGTCTACGGGGCTGAACCTGGAGGTGAAGAGGCCAGACG	3296
Qy	984	ValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArg	1003
Db	3297	CTGCGCTGTGGGACAGCTGCTCGGGAGAGACAGCATGTGTCACGTCTCTCGAGCGG	3356


```
QY 1004 LeuTyrAspProValGlnGlyGlnValLeuPheAseGlyValAspAlaLysGluLeuAse 1023
DB 3357 TTCTACGACCCCGGACGAGTGTCTTAGATGGCAAGAGTAACCACTAAAT 3416
QY 1024 ValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAseCys 1043
DB 3417 GTCCAGTGGCTCGGACCACTAGGCAATGTGTCCCAAGAGCCCACTCTGTGTACTGC 3476
QY 1044 SerIleAlaGluAseIleAlaTyrGlyAspAseSerArgValValProLeuAspGluIle 1063
DB 3477 AGCATCGCTGAGACATCGCTTATGAGACACACAGCCGGTGTGTCCAGGATGATC 3536
QY 1064 LysGluAlaAlaAseAlaAseIleHisSerPheIleGluGlyLeuProGluLysTyr 1083
DB 3537 GAGAGGCGCGCAAGGACCAACATCCACGATTCATTTAGTCACTGCTGCAATATAC 3596
QY 1084 AnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAla 1103
DB 3597 AACACAGAGTAGGACAAAGGAGCTCAGCTGTCTGGGTGGGAGAAACACGATCGCC 3656
QY 1104 IleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAla 1123
DB 3657 ATAGCAGCGCCCTGTCAGACAGCTCATATTTGCTCTGTGATGACACATCAGCC 3716
QY 1124 LeuAspAseAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArg 1143
DB 3717 CTGGATACAGAGAGTGAAAGGTTGTCCAGGAAGCTCTGGACAAAGCAGAGAGCCGC 3776
QY 1144 ThrCysLeuValValThrHisArgLeuSerAlaIleGlnAseAlaAspLeuIleValVal 1163
DB 3777 ACCTGATTTGTATGCTCTCACCCTGTCACCATCCAGAACGACAGACTTGATCGTGGT 3836
QY 1164 LeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAseArg 1183
DB 3837 ATTCAGATGGCAAGTCAAGGAGTGGCACCACAGCAGCTGTCTGGACAGAAAGGC 3896
QY 1184 IleTyrPheLysLeuValAsnAlaGlnSer 1193
DB 3897 ATCTATTTCTCCATGTCAGTGTGCAGGCT 3926
RESULT 8
AX105082 4279 bp DNA linear PAT 30-APR-2001
LOCUS Sequence 26 from Patent WO0123540.
DEFINITION AX105082
ACCESSION AX105082
VERSION AX105082.1 GI:13921232
KEYWORDS
SOURCE dog.
ORGANISM
Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE 1 (bases 1 to 4279)
AUTHORS Stocker,P.J., Steimel-Crespi,D.T., Crespi,C.L., Reif,T.C. and
Patten,C.J.
TITLE P-glycoproteins and uses thereof
JOURNAL Patent: WO 0123540-A 26 05-APR-2001;
GENTEST CORPORATION (US)
FEATURES
source
1..4279
/organism="Canis familiaris"
/db_xref="taxon:9615"
17..3862
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAC37723.1"
/db_xref="GI:13921233"
/translation="MDPEGKRGSAEKNFWMKSKSKKEKKKPTVSTFAMFRYSN
WLDRLYLVTWAAIITHGAALPLMLVFGNMTDSFANAGISRNKTPFPVINESITNT
QHFINHLEEMTYAYYSIGIGANGVLVAAYIQVSEWCLAAAGRLIKIRKQFFHIMRQ
EIGFDVHDVDELNTLDDSKINEGIDKIGMFFQSFATFTFTGIVGTFRGWLKL
VLAISPLVIGLAKWAKLSFTDKELAYAKAGAAVEALAAIRTVLAFSGGKLT
ERYNKLLEAKGIGIKATANI SIGAEFLIYAYALAFWYGTSLVLSSETYIGQVL
TVFFSLVIGAFSIGQASPIEAFAANARGAAVEIFKIIDNKPSTDSYSGSHKPDNIK
```

```
BASE COUNT 1296 a 833 c 1009 g 1141 t
ORIGIN
Alignment Scores:
Pred. No.: 3,32e-222 Length: 4279
Score: 3557.00 Matches: 683
Percent Similarity: 74.74% Conservative: 237
Best Local Similarity: 55.48% Mismatches: 269
Query Match: 59.17% Indels: 42
DB: 6 Gaps: 8
US-09-873-409-6 (1-1195) x AX105082 (1-4279)
QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
DB 167 ATGTTGGTGGGACAAATGGCTGCCATCCATGAGGCTGCATCCCTCTCTATGATGCTG 226
QY 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu----- 34
DB 227 GTTTTGGAAACATGACATAGCTTTGCAATGAGGAAATTTCAAGAAACAAACTTTT 286
QY 35 -----ValGlnThrAsnThrTyrSerPhePheArg----- 44
DB 287 CCAGTTATAATTAATGAAAGTATTACGAACAATATACACAACTTTTCATCAACCATCTGGAG 346
QY 45 -----LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe 59
DB 347 GAGGAAATGACCACGCTATGCCCTATTATTACAGTGGGATCGGTGCTGCGCTGGTGGCT 406
QY 60 GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg 79
DB 407 GCTTACATCCAGTTTCATCTGTCCTGGCCGCGAGGAGACAGATATCAAAATTAGA 466
QY 80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIle 99
DB 467 AAACAATTTTTCATGCTATCATCGACAGGAGATTGGCTGTTGACGTGATGACGTT 526
QY 100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
DB 527 GGGGAGCTTAAACACCGCGCTCAGACGATGTCTCCANAATCAATCAAGAAATGGCGAC 586
QY 119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
DB 587 AAAATGGAAATGTTCTTCAATCAATAGCAACATTTTTCACCGGTTTATAGTGGGGTTT 646
QY 139 ValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
DB 647 ACAGTGGTTTGGAGCTAACCTTGTGATTTTGGCCATCAGCCCTGCTCTTGGCATTTCA 706
QY 159 AlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
DB 707 GCCGCCATCTGGGAAAGATACTATCTTCTATTCTGATAAAGAACTCTTGGCTATGCA 766
QY 179 LysAlaGlyAlaValAlaGluValLeuLeuSerSerIleArgThrValIleAlaPheArg 198
DB 767 AAAGCTGGGACAGTAGCTGAAGAAGTCTTAGCAGCAACTCAGAACTGTGATTCCTTTGGA 826
```


QY 199 AlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 218
Db 827 GGACAAAGAAAGAACTTGAAGGTACCAACAAATTTAGAAAGCTAAAGAAATTGGG 886
QY 219 IleLysArgThrIleAlaSerLysValSerLeuGluAlaValTyrPheMetAsnGly 238
Db 887 ATAAGAAAGCTATCAGGCCAACATTTCTATTGGTGGCCGCTTTCTATTGATCTATGCA 946
QY 239 ThrTyrGlyLeuAlaPheTyrPheGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 258
Db 947 TCATATGCTCTGGCTTCTGGTATGGACCTCTTGGTCTCTCCAGTGA-----TAT 1000
QY 259 ThrIleGlyThrValLeuAlaValPheSerValIleHisSerTyrCysIleGly 278
Db 1001 ACTATTGGCAAGTACTCACTGCTCTTCTTCTGATTAAATTGGGGCTTTTAGTATTGGA 1060
QY 279 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 298
Db 1061 CAGGCATCCCAAGCAATTGAAGCATTTGCAACCGCAAGAGGAGCAGCTTATGAATCTTC 1120
QY 299 GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 318
Db 1121 AAGATAATTGACATAAACCAGCATTTGACAGCTATTGCAAGAGTGGACATAAACAGAT 1180
QY 319 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 338
Db 1181 AATATTAAAGGAAATTTGGAAATTCAAAATGTTTCACCTTCAGTTACCTTCTCGAAAGAA 1240
QY 339 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 358
Db 1241 GTTAAAGATCTTAAAGGGTCTCAACCTGAAGGTTCCAGAGTGGCGAGACAGTGGGGCTGTT 1300
QY 359 GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro 378
Db 1301 GGGAAACATGGCTCGGGGAAGAGCACCGTCGAGCTGATGAGAGGCTCTATGACCCC 1360
QY 379 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 398
Db 1361 ACAGATGGCATGGTCTGATTGATGGACAGGACATTAGGACCATAAATGTAAGGCATCTT 1420
QY 399 ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrIleSerAsn 418
Db 1421 CGGAAATTAATCTGGTGTGGTGAAGTCAAGAGCTGTGTGTTGTCACACAGCAGTAGCTGAA 1480
QY 419 AsnIleLysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGlu 438
Db 1481 AACATTCGATGGCCCGGAAAAATGTCCACATGATGAGATTGAGAAAGCTGTTAAGGAA 1540
QY 439 AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 458
Db 1541 GCCATGCTTATGATTTTATCATGAACCTACCTAATAAATTTGACACTCTGGTTGGAGAG 1600
QY 459 LysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal 478
Db 1601 AGAGGGCCAGCTGAGTGGTGGAGCAAAACAGAGAATCGCCATTGCTCGGGCCCTGGTT 1660
QY 479 ArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 498
Db 1661 CGCAACCCCAAGATTTCTCTGATGGAGCAACAGCTCAGCTCTGGACATCGAAAGTGA 1720
QY 499 SerAlaValGlnAlaLeuGluLysAlaSerLysGlyArgThrIleValValAla 518
Db 1721 GCAGTGGTTTCAAGTGGCCCTGGATGAAGCCAGAAAGCCGGACTACCATTTGTGATAGCT 1780
QY 519 HisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeu 538
Db 1781 CATCGTTTGTCTACAGTTGTAATGTCGATGTCATTCGTTGTTGATGATGAGTCAAT 1840
QY 539 AlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrSerLeuVal 558
Db 1841 GTGGAGAAAGGAATCATGATGAACCTCATGAAGAGAGGCAATTTACTTCAAACTTGTC 1900
QY 559 MetSerGlnAspIleLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu 578

Db 1901 ACAATGCAG---ACAAGAGGAATGAAATTTAGTTAGAAAATGCCACTGGTGAATCCAAA 1957
QY 579 ArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle--- 597
Db 1958 AGTGAAGTGAATGCCTTGGAAATG---TCTCCAAAAGATTCAGGGTCCAGTTTAATAAAA 2014
QY 598 -----AspLysAlaGluGluSer 603
Db 2015 AGAAGATCAACTCGCAGGAGTATACATGCACCACAGGCCAACAGACAGAAAGCTTGGTACA 2074
QY 604 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 623
Db 2075 AAAGAGGACTTGAATGAGAAATGATCCCTCCAGTTCTCTCGAGGATTCGAAGCTGAAC 2134
QY 624 LysProGluTyrProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 643
Db 2135 TCACCTGAATGGCTTATTTTGGTGGTATATATTTTGGCTATATAAACGGAGGCCCTG 2194
QY 644 HisProValPheSerIleIlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAsp 662
Db 2195 CAACACGACATTTCAATAATATTTCAAGGATATATAGGATCTTTTACCCGAGATGAGAT 2254
QY 663 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 682
Db 2255 CCTGAAACAAAACGACAGAAATAGTAACTGTTTCTGTATTTGTTCTAGTCTCTGGAAT 2314
QY 683 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 702
Db 2315 ATTTCTTTTATACATTTTCTCCAGGGCTTCACATTTGGCAAGCTGGGGAGATCCCT 2374
QY 703 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPhe 722
Db 2375 ACTAAGCGCTTCGATACATGTTTTCAGATCCATGCTGAGACAGGATGTCAGCTGGTT 2434
QY 723 AspGluLysGluAsnSerThrGlyGlyLeuThrIleLeuAlaIleAspIleAlaGln 742
Db 2435 GATGACCCATAAAACACCACTGGAGCATTTGACAAACAGGCTTGCCCAATGATGCGGCTCA 2494
QY 743 IleGlnGlyAlaThrGlySerIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 762
Db 2495 GTTAAAGGGCTATAGTTCCAGGCTTGCTGCTATACCAGAAATATAGCAATCTTGGG 2554
QY 763 LeuSerValIleIleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIle 782
Db 2555 ACAGGCAATTATATATCTTAATCTATGTTGGCAATTAACACTTTTACTCTTAGCAATT 2614
QY 783 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 802
Db 2615 GTACCCATCATTCGAATAGCAGGAGTGTGTAATGAAATGAAATGTTGTCTGGACACAGCTG 2674
QY 803 LysAspLysGlnGluLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 822
Db 2675 AAAGATAAGAAAGAGCTAGAAGGCTGGGAAGATTGTACAGAAGCCATCGAAACTTC 2734
QY 823 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 842
Db 2735 CGAACCTGTTGTTTCTTGTGCTCGGAGCAGAGATTGTAATACATGATGCACAGAGTTTG 2794
QY 843 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 862
Db 2795 CAAGTACCATACAGAAACTCTTCTGAGGAAAGCACACATCTTCGGGGTCTCATTTTCTATC 2854
QY 863 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 882
Db 2855 ACCCAGGCATGATGATTTTCTCTATGCTGGCTGTTTCCGGTTTGGTGGCTACTTGGTG 2914
QY 883 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 902
Db 2915 GCAATGAGTTCAAGAACTTTCAGGATGTTCTTTTGGTATTTCTCAGCTATTGCTTGGT 2974
QY 903 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 922

Db	2975	GCATCGCAGTGGGCGAGTCAGTTCATTGCTCCTGACTATGCCAAGCAAGTATCA	3034
Qy	923	AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu	942
Db	3035	GCAGCCACGTCATCATGATCATTTGAATAAAGCCCTCTGATTGACAGCTACAGCCCTCAC	3094
Qy	943	GlyLysLysProAspThrCysGluGluGluAsnLeuGluPheArgGluValSerPhePheTyr	962
Db	3095	GGCCTCAAGCCAAATACGTTGGGAAGGAATGTGACATTTTAATGAGGTCGTGTTCAACTAT	3154
Qy	963	ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys	982
Db	3155	CCCACTGACACACATCCCCGTCCTCAGGCGCTGAGCCTCGAGTGAAGAGGCCGAC	3214
Qy	983	ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln	1002
Db	3215	ACGCTGCCCTCGTAGTAGCAGTGGCTGTGGGAAGACACAGTGTTCAGCTCCTAGAG	3274
Qy	1003	ArgLeuTyrAspProValGlnGlnValLeuPheAspGlyValAspAlaLysGluLeu	1022
Db	3275	CGCTTCATGACCCCTTGGCTGCTTTCAGTCTAATTCATGGCAAGAGATAAAGCACCTG	3334
Qy	1023	AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn	1042
Db	3335	AATGTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCTCAGGAGCCCATCCTGTTTGC	3394
Qy	1043	CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValProLeuAspGlu	1062
Db	3395	TGCAGCATTCGCCGAGAACATTCGCTATGGAGACACAGCGCGGTCTGATCACATGAAG	3454
Qy	1063	IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys	1082
Db	3455	ATTGTCAGCAGCAGCCAAAGGAGGCGCAACATACACCACTTCATCGAGACACTCCCTG	3514
Qy	1083	TyrAsnThrClnValGlyLeuLysGlyAlaGlnLeuSerClyGlyGlnLysGlnArgLeu	1102
Db	3515	TACAAACAGCAGTAGGAGAGCAAGGAACCCAGCTCTCTGGTGGCCAGAAACAGCGCAT	3574
Qy	1103	AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer	1122
Db	3575	GCATAGCTCGCGCTCTGTAGACAGCCTCATATTTGCTTTGGATGAGTACATCA	3634
Qy	1123	AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly	1142
Db	3635	GCTCTGGATACAGAAAGTGAAGAGTTGTCCAAGAGCCCTCGACAAAGCCAGAGAAGGC	3694
Qy	1143	ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal	1162
Db	3695	CGCACCTGCTATTGTGATCGCCACCGCTTGTCCACCATCCAGAAATGAGATTAAATAGT	3754
Qy	1163	ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg	1182
Db	3755	GTGTTTCAGATGGCAAGTCAAGAGCATGGCACATCCACATCCAGCTGTGCGCCAGAAA	3814
Qy	1183	AspIleTyrPheLysLeuValAsnAlaGlnSer	1193
Db	3815	GGCATCTATTTTCCATGGTCCAGTCCAGGCT	3847
RESULT	9		
LOCUS	AX108656	4195 bp	DNA
DEFINITION	Sequence 3 from Patent WO0123565.		
ACCESSION	AX108656		
VERSION	AX108656.1	GI:13923888	
KEYWORDS	crab-eating macaque.		
SOURCE	Macaca fascicularis		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;		
	Cercopitheciinae; Macaca.		
REFERENCE	1 (bases 1 to 4195)		
AUTHORS	Stocker,P.J., Steimel-Crespi,D.T. and Crespi,C.I.		
TITLE	P-glycoproteins from macaca fascicularis and uses thereof		

JOURNAL	Patent: WO 0123565-A 3 05-APR-2001;		
FEATURES	GENTEST CORPORATION (US)		
source	Location/Qualifiers		
	1..4195		
	/organism="Macaca fascicularis"		
	/db_xref="taxon:9541"		
CDS	100..3951		
	/note="unnamed protein product"		
	/codon_start=1		
	/protein_id="CAC37765.1"		
	/db_xref="GI:13923889"		
	/translation="MDLEGDRNGGAEKKNFFKLNKSKKDKKERKPTVSVFMSFRYSN		
	WLDKLYMVVGTAAIIGHGAGLPLMLVFGDMTDFANAGNLGDGLALLPNTNNSNIT		
	DTVPNMLEEDMTRYIYSGIGAGLVAAVYQVSWFLAAGRQHKIRKQPFPHAIR		
	QEIFGDFVHDVGEMLNTRLTDVSGINEGIDKIGMFQSMATFFTFITVFGFRWKL		
	LVLIAISPLGLSAAWAKILSFTDKELLAYAKAGAAVEALVIAIRTVIAFGQKKE		
	LERYKNLEAKRIGIKKAITANISIGAFLLIYASALAPWYGTVLVLVSKEYSIGOV		
	LTVPSVLIGAFSVGOAAPSIEAFANARGAPEFKIIDNKPSIDSYSKSHKPNIK		
	GNLFRNVHPSYPRKSEVKILKGLNLKQSGQTVALVNSGCGKSTTVQLMQRLDYPT		
	EGMSVDGQDITINVRFLREIIGVVSQEPVLFIATINIRYGRHEDTMDIEKAVK		
	ENAYDFIMKLPQKFDTLVGERGQLSGGQKQRIARALVRNPKILLIDLEATSLDT		
	ESEAIVQALDKARKGRTTIVIAHRLSTVRNADVIAGFDGVIKRGHDELKKEKGI		
	YKLVMTOTAGNELENADESKSEIDTLEMSHSDSGSLIRKSTRRSVRGSOQD		
	RKLSKEALDESIIPVSWRIMKLNLTMPYFVVGFCALINGLOPAFAVFSKIIIG		
	IFTRNDRAETKRONSLFLLVLGIVSFITFFLOGTFPGKAGBILKRLRYMWFPS		
	MLRQDSWFDPPKNTTGALTTRLANDAQVKAIGSLRAIITQNTIANIGTGIIISLY		
	QMOTLLFLLAIVPIIAIVGVEMKMLSQALKDKKELEGAKIATIEINFRVVSLLT		
	QEQFPHMYDOSLVQVYRNSLRKAHIFGITFTQAMMYFVAGCFRFGAYLVASHLM		
	SFEDVLIVFSAVFGAMAVGVSSPAPDYAKAVSAANIIMTIEKTPIDSYSTGLK		
	PNTLEGNVTFNEVFNYPTRLDIPVLQSLSEVKKGTALVLVSGSGCKSTVQLIER		
	FYDLAAGKVLIDGKEIKQNLVQWLRHAGIIVSQEPILFDCSISENIAIYGNRSRVSQE		
	EIVRAAKEANHAFTIESLPNKYSTRVGDKTQLSGQKQRIARALVRNPKILLIDLE		
	ATSALDTSEKVVQALDKAREGRTCTIVIAHRLSTIQNALDIIVFQNGRVEKHGTHQ		
	LLAOKGIVFMSVQAGAKRO"		
BASE COUNT	1230 a 801 c 1039 g 1125 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	3,49e-222	Length:	4195
Score:	3556.50	Matches:	680
Percent Similarity:	75.16%	Conservative:	246
Best Local Similarity:	55.13%	Mismatches:	263
Query Match:	59.16%	Indels:	43
DB:	6	Gaps:	9
US-09-873-409-6 (1-1195) x AX108656 (1-4195)			
Qy	1	MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu	20
Db	250	ATGGTGGTGGGAACCTTTGGCTGCCATCATCCATCGAGCTGGACTTCCTCTCATGATGCTG	309
Qy	21	ValLeuGlyGluMetSerAspAsnLeuIleSer-----GlyCys	33
Db	310	GTGTTTGGAGACATGACGGATACCTTTGGCAATGACGAGAAATTTAGGAGATTAGGAGCT	369
Qy	34	LeuValGlnThrAsnThrTyrSer-----GlyCys	41
Db	370	CTGTTGTTTAAACAACACTAATAGCAGTAATATCAGTATACAGTGCCTGATGAATCTG	429
Qy	42	-----PhePheArgLeuThrLeuTyrTyrValGlyIleGlyValalaalaLeuIle	58
Db	430	GAGGAAGATATACACCAAGGATGCTTATTATTACAGTGGAAATTTGGTGGTGGTGGT	489
Qy	59	PheGlyTyrIleGlnIleSerLeuTyrIleIleThrAlaAlaArgGlnThrLysArgIle	78
Db	490	CTGCTTACATTCAGGTTTCATTTTGGTGGCAGCTGGGAAGAACATAACACAAAT	549
Qy	79	ArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAsp	98
Db	550	AGAAACAGTTTTTTCATGCTATAATGCACAGACAGATAGGCTGGTTTGTATGTCAGAT	609
Qy	99	IleGlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGly	117

Db 610 GTTGGGAGCTTAACACCCGGCTTACAGATGATGTCTCAAGATTAAATGAAGAAATTGGT 669
Qy 118 AspIysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGly 137
Db 670 GACAAAATTGGAATGTTCTTTCAGTCAATGGCAACATTTTTCATCTGGTATTAGTAGGA 729
Qy 138 LeuValIysGlyTrpIysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAla 157
Db 730 TTTACAGCTGGTGGAGCTAACCCCTGTGATTTGGCCATCATCTCTGTCTTGGACTG 789
Qy 158 SerAlaAlaCysSerArgMetValIleSerLeuThrSerIysGlnLeuSerAlaTyr 177
Db 790 TCAGCTGCAGCTGGGCAAGATACCTCTTCATTACTGATAAGAACTCTTAGCTTAT 849
Qy 178 SerIysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPhe 197
Db 850 GCAAAAGCTGGAGCAGTAGCTGAAGAGGCTTGTGGCAATTAGAACTGTGATGGATTT 909
Qy 198 ArgAlaGlnGluIysGluLeuGlnArgTyrThrGlnAsnLeuIysAspAlaIysAspPhe 217
Db 910 GGAGGACAAAGAAAGAACTCGAAAGGTCAACAACAAAATTTAGAAAGACTAAAGAAATT 969
Qy 218 GlyIleIysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsn 237
Db 970 GGGATAAAGAAAGCTATTACAGCCCAATATTCTATAGGTGCTTTCCTGTATATCTAT 1029
Qy 238 GlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsnGlyGluProGly 257
Db 1030 GCATCTTATGCTCTGGCCCTCTGTGTATGGGACCACTTGTCTCTCAAGGAA----- 1083
Qy 258 TyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerTyrCysIle 277
Db 1084 TATCTATTGGACAAGTACTCAGCTGATTTCTTTCGTATTAAATGGGGCTTTTAGTGTT 1143
Qy 278 GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 297
Db 1144 GGACAGGCATCTCAAGCAATTGAAGCATTTGCAAAATGCAAGAGGAGCAGCTTTGAAATC 1203
Qy 298 PheGlnValIleAspIysIysProSerIleAspAsnPheSerThrAlaGlyTyrLysPro 317
Db 1204 TTCAGATAATTGATAATAAGCCCAAGTATTGACAGCTATTGCAAGAGTGGGCGCAACCA 1263
Qy 318 GluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgPro 337
Db 1264 GATAAATTAAAGGAATTTGGAATTCAGAAATGTTCACTTCAGTACCCATCTCGAAA 1323
Qy 338 SerIleIysIleLeuIysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu 357
Db 1324 GAAGTTAAGATCTTGAAGGGCTTGAACCTGAAGGTGCAGAGTGGCGACGCTGGCCCTG 1383
Qy 358 ValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAsp 377
Db 1384 GTTGGAAACAGCGCTGTGGGAAGAGCACACCGTTCAGCTGATGCAGAGCTTTATGAC 1443
Qy 378 ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis 397
Db 1444 CCCACAGGGCATGGTCAGTGTGATGACAGCAGATATTAGGACCAATAAGAAAGTTT 1503
Qy 398 TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSer 417
Db 1504 CTACGGGAAATCATCGGTGTGTGAGTCAAGCAACCTGATTTGTTTGGCCACCAATAGCT 1563
Qy 418 AsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArg 437
Db 1564 GAAACCATTCCTATGGTCGTGAGAGATGTCCACATGATGATGATGATGAGAAAGCTGTCAAG 1623
Qy 438 GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly 457
Db 1624 GAAGCCATATGCTATGACTTTATCATGAACCTGCCTCAGAAATTTGACACCTGTTTGA 1683
Qy 458 GluIysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu 477
Db 1684 CAGAGAGGGGCCAGCTGAGTGGTGGGCGAGAGCAGAGGATCGCCATTGGCAGCTGCCCTG 1743

Qy 478 ValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 497
Db 1744 GTTGCACCAACCCCAAGATCTCTCTGTGACGAGGCCACGTCAGCCTTGGACACAGAACT 1803
Qy 498 LysSerAlaValGlnAlaLeuGluIysAlaSerLysGlyArgThrThrIleValVal 517
Db 1804 GAAGCAGTGGTTCAGGTGGCTCTGATTAAGGCCAGAAAAGGTCGGACCACTTGTGATA 1863
Qy 518 AlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMet 537
Db 1864 GCTCATCGTTTCTACGGTTCTGTAATGCCGACGTCATCGCTGGTTTCGATGATGGAGTC 1923
Qy 538 LeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeu 557
Db 1924 ATTGTGGAGAAAGAAATCATGATGAGCTCATGAAAGAGAAAGGCATTTACTTCAACTT 1983
Qy 558 ValMetSerGlnAspIle-----LysLysAlaAspGlu----- 568
Db 1984 GTCACAAATCGACACAGCAGGAAATGAAATTTGAATTAGAAAATGCAGCTGATGAATCCAAA 2043
Qy 569 ---GlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeu----- 584
Db 2044 AGTGAATTTGATACCTTGGAAATGCTTTCACATGATTCCAGGATCCAGTCTAATAAGAAAA 2103
Qy 585 -----ProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGlu 602
Db 2104 AGATCCACTCGTAGGAGTGTCCGTGGATCAAGGCCAA-----GACAGAAAGCTTAGT 2157
Qy 603 SerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeu 622
Db 2158 ACCAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGTGA 2217
Qy 623 AsnLysProGluTyrProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThr 642
Db 2218 AATTAACTGAGTGGCTTATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2277
Qy 643 ValHisProValPheSerIleIlePheAlaLysIleIleThrMetPhe---GlyAsnAsn 661
Db 2278 CTGCAACCAAGCATTTGCAGTAAATATTTTCAAAGATTATAGGGATTTTACAAGAAATGAT 2337
Qy 662 AspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGly 681
Db 2338 GATCGCGCAACAAAACGACAGAAATAGTAATCTTTTCACTATTGTTCTTAGTCTCTGGA 2397
Qy 682 ValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIle 701
Db 2398 ATGTGTTCTTTTATTAATATTTTCTTCAAGGCTTCACATTTGGCAAGAGCTGAGAGATC 2457
Qy 702 LeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyr 721
Db 2458 CTCACCAAGCGGCTCCGATACATGTTTCCGATCCATGCTCAGACAGGATGTGAGCTGG 2517
Qy 722 PheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAla 741
Db 2518 TTTGATGACCTTAAAAACACCACTGGAGCATTGACTACAGGCTCGCCCAATGATGTGCT 2577
Qy 742 GlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMet 761
Db 2578 CAAGTTAAAGGGCTATAGTTCAGGCTTGCTATTAATTTACCAGATATAGCAATCTT 2637
Qy 762 GlyLeuSerValIleIleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSer 781
Db 2638 GGCACAGGAATAATTATATCTTAAATCTATGTTGGCAACTGACACTGTACTCTTAGCA 2697
Qy 782 IleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAla 801
Db 2698 ATTTGATCCCATCTATGCAATAGCAGGAGTGTGTAATGAAATGTTGTTGTGACCAAGCA 2757
Qy 802 AsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsn 821
Db 2758 CTGAAAGATGAAGAAACCTAGAAGGTGCTGGGAGAGATCGCTACTGAGCAATAGAAAAC 2817

Alignment Scores:

Pred. No.: 4,98e-222 Length: 4927
Score: 3555.50 Matches: 682
Percent Similarity: 75.78% Conservative: 244
Best Local Similarity: 55.81% Mismatches: 263
Query Match: 59.14% Indels: 33
DB: 10 Gaps: 9

US-09-873-409-6 (1-1195) x AF257746 (1-4927)

```
QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 233 ATGCTGCTGGAACTCTGGGGGCGCATATCATCGAATTCGCTCCACATTATGCTG 292
QY 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
Db 293 GTCTTTGGAGACATGACAGATAGCTTTGGCAAAAT-----GTAGAAACAACCGTAGTATG 346
QY 41 SerPhePheArgLeuThr-----GTAGAAACAACCGTAGTATG 46
Db 347 AGTTTCTACATGCTACAGACATATATGCCAAGCTGGAGGACGAAATGACCGTAGCGCC 406
QY 47 LeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeu 66
Db 407 TACTATTACACGGGCGCATTTGGTGGCGGTGCTCATCGCTTACATCCAGGTTTCACTT 466
QY 67 TrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGlnPhePheHisSerVal 86
Db 467 TGTGCTGCGCAGCTGGGAGACAAATACACAAGATTAGGACGAAGTTTTTCCATGCGCATC 526
QY 87 LeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGluLeuAsnThrArgMet 106
Db 527 ATGAACTCAGAGATAGCTGGTTTGACGTGCATGACGTGGGAGCTCAACACCGGCTC 586
QY 107 Thr----AspIleAspIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGln 125
Db 587 ACAGATGACGTCTCCAAAATTAATAGAGGAATTGGTGACAAAATTTGGAATGTTCTTTTCAG 646
QY 126 AsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLysGlyTyrLysLeuThr 145
Db 647 GCAATGGCAACATTTTGTGGTGTATATATAGGATTTACTCGCGCTGGAAAGCTTAAC 706
QY 146 LeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAlaAlaCysSerArgMet 165
Db 707 CTGTGTGATTTGGCCATCAGCCCTGTCTTGGACTGTGACGTGTATTTGGGCAAGATA 766
QY 166 ValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGlu 185
Db 767 TTGTCTTCATTTACTGATAAGAACTCCAGGCTTATGCAAAAGCTGGAGCAGTTGCTGAA 826
QY 186 GluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGln 205
Db 827 GAAGTCTTAGCAGCCATCAGAACTGTGATTTGCCCTTTGGAGGACAAAAGAACTTTGAA 886
QY 206 ArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSer 225
Db 887 AGGTACAAATAACAAATTTGGAGAAGCTTAAAGGCTTGGGATAAAGAAGCTATCACGGCC 946
QY 226 LysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTyr 245
Db 947 AACATTTCCATGGGTGCAGCTTTTCTGTTATCTATCATCATATGCTCTGGCAATCTCG 1006
QY 246 TyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAla 265
Db 1007 TATGGGACTTCTTGGTCACTCAAAAGAA-----TACACTATTGGACAAGTGTCTCACT 1060
QY 266 ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGlu 285
Db 1061 GTCTTTTCTGTATTAAATGGAGCATTCAGTGTGGCGGCTCTCCAAATATTGAA 1120
QY 286 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysPro 305
Db 1121 GCCTTCGCAATGCTAGAGGAGCAGCTTATGAAGTCTTCAGTATAATTTGATAATAAGCCC 1180
```

```
QY 306 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 325
Db 1181 AGTATAGACAGCTTCTCAAGAGTGGGCACAAACCGACACATACAGGAAATTTGGAA 1240
QY 326 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu 345
Db 1241 TTCAAAATATTTCACTTCACTTACCGTCTCGAAAAGAGCGTTTCAGATCTTTGAAGGCGCTC 1300
QY 346 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 365
Db 1301 AACCTGAAGGTGAAGAGCGGCGAGCGTAGCGTGTGGCAACAGTCGCTGTGGAAA 1360
QY 366 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 385
Db 1361 AGCACAACTGTCAGCTGCTGCAGAGGCTCTACGACCCCTAGAGGGCGGAGTCACTATC 1420
QY 386 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal 405
Db 1421 GACGGACAGACATCAGGACCATCAATGTGAGTATCTCGGGAATCATTTGGGTGGTG 1480
QY 406 SerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAsp 425
Db 1481 AGTCAGAAACCGGTGCTGTTCACCAATTTGCCAATTTGCCGAAACATTCGCTATGCCGAGAA 1540
QY 426 AspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIle 445
Db 1541 AACTCACCATGATGATAGAGAAAGTGTCAAGAGAGCCATTCCTATGNTTCAATC 1600
QY 446 MetGluPheProAsnLysPheAsnThrLeuValGlyGlyGlyAlaGlnMetSerGly 465
Db 1601 ATGAAACTGCCCCACAAATTTGACACCTTGTGGTGAGAGGGCGGAGCTGAGTGGG 1660
QY 466 GlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIle 485
Db 1661 GGACAGAAACAGAGGATCGCCATTTGCCCGGCGCTGTGTCGCAACCCCAAGATCTTTTG 1720
QY 486 LeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeu 505
Db 1721 TTGATAGAGCCACGTCAGCTTGGACACAGAAAGCGAGCGGTTCAGGCGGCTCTG 1780
QY 506 GluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIleArg 525
Db 1781 GATAAGCTAGAGAGGCGGACCCACCAATTTGTATAGCTCACCGCTTGTCTACAGTTCG 1840
QY 526 SerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAla 545
Db 1841 AATGCTGACGCTCATCTGCTGTTTGTATGTTGTGTTGTCATTGTGGAGCAAGGAATCATGAT 1900
QY 546 GluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln-----MetThrTyr 561
Db 1901 GAGCTCATGAGAGAAAGAAAGAAATTTACTCAAACTTGTATGACTCAGACAGCAGGAAAT 1960
QY 562 AspIleLysLysAlaAspGluGlnMetGluSer-----MetThrTyr 575
Db 1961 GAAATTTGAATTAGGAATGAAGCTTGTGAATCTAAAGATGGAATTCATAATGGGACATG 2020
QY 576 SerThrGluArgLysThrAsnSerLeu-----ProLeuHisSerValLysSerIleLys 593
Db 2021 TCTTCAAAAGATTTCAGGATCCAGTCTTAATAAGAAGAAGATCAACTCGCAAAAGCATCCGT 2080
QY 594 SerAspPheIleAspLysAlaGluGluSerThr---GlnSerLysGluIleSerLeuPro 612
Db 2081 GGGCCCATGATCAAGACGGGGAACCTAGCACCAAGAGGCTCTGGATGACGACGTACCT 2140
QY 613 GluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTyrProPheValValLeu 632
Db 2141 CCAGCTTCTTTTGGCGGATCCTGAAGTTGAATTCAACTGAATGGCTTATTTTTGGTT 2200
QY 633 GlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePheAla 652
Db 2201 GGTGTATTTTGTGCCATAATAATGAGGCTTCGACACCGACGATCTCTCCATAATATTTTCA 2260
```


Db	3341	GTGTTTCAGATGCGCAAGAAATAAAGCAACTCAATGTCACATGGCTCGCGGCCCACTG	3400
Qy	1032	AlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyr	1051
Db	3401	GGCATGTGTGCCAGAGGCCATCTCTGTTTGACTGCAGATCGCCGAGAACATTTGCCTAC	3460
Qy	1052	GlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAAsn	1071
Db	3461	GGAGACAAACAGCGTGTCTCATGAGGAGATCGTGAAGGCAGCCCAAGGAGCCCAAC	3520
Qy	1072	IleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGly	1091
Db	3521	ATCCACCAAGTTTCATCACTCACTGCGTGAAGAATAACAACACGAGAGTGGGACAAAGGG	3580
Qy	1092	AlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaAtrqAlaLeuGlnLys	1111
Db	3581	ACTCAGCTGTGGGGCGGAGAGAGAGCGATCGCGCGCCCTCGTCAGACAG	3640
Qy	1112	ProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysVal	1131
Db	3641	CCTCACATCTTACTTCTGATGAAGCAGCATCAGCTCTGGATACGAGAGTGAAGAAGTC	3700
Qy	1132	ValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArg	1151
Db	3701	GTCAGGAAGCGGTGACAAAGCCAGGGAAGCCGACCTCGCATTTGTGATCGCACCGC	3760
Qy	1152	LeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGlu	1171
Db	3761	CTGTCCACCATCCAGAACCGAGACTTGATCTGTGTGATTCAGACCGCGCGTCAGGAG	3820
Qy	1172	GlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAla	1191
Db	3821	CACGGCACCCACAGCAGCTGTGTGGCCAGAAAGCATCTATTCTCGATGTCAGTGTG	3880
Qy	1192	GlnSer 1193	
Db	3881	CAGGCT 3886	
RESULT 11			
LOCUS	AX322787	3860 bp	DNA
DEFINITION	Sequence 1 from Patent WO0192877.		linear
ACCESSION	AX322787		
VERSION	AX322787.1	GI:18093766	
KEYWORDS	human.		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1		
AUTHORS	Sorrentino, B. and Schuetz, J.		
TITLE	Method of identifying and/or isolating stem cells		
JOURNAL	Patent: WO 0192877-A 1 06-DEC-2001;		
ST. JUDE	CHILDREN'S RESEARCH HOSPITAL (US)		
FEATURES	Location/Qualifiers		
source	1..3860		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
BASE COUNT	1135 a 746 c 957 g 1022 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	4..56e-222	Length:	3860
Score:	/ 3554.00	Matches:	682
Percent Similarity:	75.63%	Conservative:	246
Best Local Similarity:	55.58%	Mismatches:	263
Query Match:	59.12%	Indels:	36
DB:	6	Gaps:	9
US-09-873-409-6 (1-1195) x AX322787 (1-3860)			
Qy	1	MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu	20

[illegible]

D	b	1225	ATCTTGAAAGCCCTGAACCTGAAAGTGCAGAGTGGGCAGACGGTGGCCCTGGTGTGGAAC	128
Q	y	361	AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyArgProAspAsp	380
D	b	1285	AGTGGCTGTGGGAAGACACAACAGTCCAGCTGTATGCAGAGGCTCTATGACCCACAGAC	1344
Q	y	381	GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyArgAsp	400
D	b	1345	GGCATGTCTAGCTGTGATGCAGCAGGATATTAGGACCATAAATGTAAAGTTTCTACGGGA	1404
Q	y	401	HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle	420
D	b	1405	ATCAATTCGTGTGTGAGTCAGGAACCTGTATTGTTGGCCACACAGATAGCTGAAAAACAT	1464
Q	y	421	LysTyArgArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn	440
D	b	1465	CGCTATGGCCGTGAANAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGNAGCCNAT	1524
Q	y	441	AlaTyArgPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly	460
D	b	1525	GCCTATGACTTTATCATGAACCTGCCTCATAAATTTGACACCTGTTGGAGAGAGGG	1584
Q	y	461	AlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn	480
D	b	1585	GCCCAGTTGAGTGGTGGCAGACAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC	1644
Q	y	481	ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla	500
D	b	1645	CCCAAGATCTCTCTGTGATGAGGCCAGCTCAGCTCTGGACACAGAAAGCGACGATG	1704
Q	y	501	ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg	520
D	b	1705	GTTCCAGTGGCTCTGGATAAGGCCAGAAAGGTCGACCCACCATTTGTATAGTTCATCGT	1764
Q	y	521	LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu	540
D	b	1765	TGTCTACAGTTCGTAAATGCTGACGTCACTCGCTGGTTTCGATGATGGAGTCAITGTGGAG	1824
Q	y	541	LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyTySerSerLeuValMetSer	560
D	b	1825	AAAGAAATCATGATGAACCTCATGAAGAGAAGGCAATTTACTTCCAACTTGTCACATG	1884
Q	y	561	Gln-----AspIleLysLysAlaAspGluGlnMetGluSerMetThr	574
D	b	1885	CAGACAGCAAGAAATGAAGTTGAATAGAAAAATGCAGCTGATGAATCCAAAAGTGAAATT	1944
Q	y	575	TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal-----	589
D	b	1945	GATCCCTTGGAAAATGTCTTCAAAATGATCAAGATCCAGTCTAATAAGAAAAAGATCAACT	2004
Q	y	590	--LysSerIleLys--SerAspPheIleAspLysAlaGluLysThrGlnSerLys	607
D	b	2005	CGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACCAAGAGGCTCTG	2064
Q	y	608	GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrp	627
D	b	2065	GATGAAGATATACCTCAGTTTCCTTTGGAGATTATGAAGCTAAATTTAACTGAATGG	2124
Q	y	628	ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe	647
D	b	2125	CCTTATTTTGTGTGTGTATTGTCGCAATATATAAATGGAGGCTGTCAACACCAATTT	2184
Q	y	648	SerIleIlePheAlaLysIleIleThrMetPheGlyAsn--AsnAspLysThrThrLeu	666
D	b	2185	GCAATAATATTTTCAAGATATTAGGGGTTTTTACAAAGAAATTTGATGATCCTGAAACAAA	2244
Q	y	667	LysHisAspAlaGluIleTySerMetIlePheValIleLeuGlyValIleCysPheVal	686
D	b	2245	CGACAGATAGTAACCTGTTTTTCACTATTGTTCTAGCCCTTGGAAATATTCTTTTATT	2304
Q	y	687	SerTyPheMetGlnGlyLeuPheTyGlyArgAlaGlyGluIleLeuThrMetArgLeu	706
D	b	2305	ACATTTTTCCTTCAAGTTTCAATTTGGCAAAGCTGGAGAGATCTCACCAAGCGGCTC	2366


```
QY 707 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 726
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2365 CGATACATGTTTTCGATCCATGCTCAGACAGGATGGAGTTGGTTGATGACCTAAA 2424

QY 727 AsnSerThrGlyGlyLeuThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 746
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2425 AACACCACTGGAGATTGATACAGGCTCGCAATGATGCTCAAGTTAAAGGGGCT 2484

QY 747 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 766
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2485 ATAGGTTCCAGGCTTGCTGTAATTACCCAGATATAGCAATCTTCGGACAGGAATAATT 2544

QY 767 IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 786
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2545 ATATCTCTCATCTATGTTGGCAACTAACACTGTTACTCTTAGCAATTGTACCCATCATT 2604

QY 787 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 806
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2605 GCAATAGCAGGAGTTGTTGAAATGAAATGTTGTCGGACAAGCACTGGAAGATAAGAAA 2664

QY 807 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 826
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2665 GAACTAGAGGTCCTGGAGAGATCGTACTGAAGCAATAGAAACTTCCGAAACGGTTGTT 2724

QY 827 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 846
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2725 TCCTTGACTCAGGACGAGAAAGTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATAC 2784

QY 847 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 866
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2785 AGAAACTCTTTGAGGAAAGCACACATCTTTGGAAATTACATTTTCTTCACCCAGGCAATG 2844

QY 867 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 886
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2845 ATGTATTTTCTATGCTGATGTTCCGGTTTGGAGCCTACTTGTGGGCACATAAATC 2904

QY 887 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 906
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2905 ATGAGCTTTCAGGATGTTCTGTAGTATTTCAGCTGTTGCTTGTGGTCATGGCGCGTG 2964

QY 907 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 926
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2965 GGCAAGTCAGTTCATTGTCTCTGACTATGCCAAGGCCAAATATATCAGACGCCACATC 3024

QY 927 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 946
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3025 ATCATGATCATGAAAAACCCCTTTGATTGACAGCTACAGCAGGAGCCCTAATGCCG 3084

QY 947 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 966
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3085 AACACATTGGAAGGAATGTACATTTGGTGAAGTTGATTCAACTATATCCACCCGACCG 3144

QY 967 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 986
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3145 GACATCCAGCTGCTTCAGGAGCTGAGCTCGAGGTGAAGAGGCCAGACGCTGGCTCTG 3204

QY 987 ValGlySerSerGlyLysLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 1006
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3205 GTGGCGACAGTGGCTGTGGAGAGACACAGTGTCTCAGCTCTCTGGAGCGGTTCTACGAC 3264

QY 1007 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 1026
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3265 CCCTTGGCAGGAAAGTGCTGCTTGATGGCAAGAAATAAGCGACTGAATGTTCACTGG 3324

QY 1027 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 1046
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3325 CTCGACACACCTGGGCATCGTGTCCAGGAGGCCATCTCTGTTTGCATCGACATTCGT 3384

QY 1047 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 1066
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3385 GAGAACATTCCCTATGTGAGACAAACAGCCGGTGTGTACAGGAAGAGATCTGTGAGGGCA 3444
```

```
QY 1067 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 1086
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3445 GCAAGAGAGGCCCAACATACATGCTTTCATCGATCTACTGCTTAATAATATAGACATA 3504

QY 1087 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 1106
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3505 GTAGGAGACAAAGGAACCTCAGCTCTCTGTGGCCAGAAACACGCATTGCCATAGCTCGT 3564

QY 1107 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 1126
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3565 GCCTTGTGTAGACAGCTCATATTTTGTCTTTGGATGAAGCCACGTCAGCTCTGGATACA 3624

QY 1127 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1146
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3625 GAAAGTGAAGGGTTGTCGAAGAAGCCCTGGCAAAAGCCAGAGAAGCCGACCTGCATT 3684

QY 1147 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1166
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3685 GTGATTGCTCACCGCTGTCACCATCCAGATGCAAGACTTAATAGTGTGTTTCAGAAAT 3744

QY 1167 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1186
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3745 GGCAGAGTCAAGGAGCATGCGCAGCATCAGCAGCTCTGCGACAGAAAGCATCTATTIT 3804

QY 1187 LysLeuValAsnAlaGlnSer 1193
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3805 TCAATGTCAGTGTCCAGGCT 3825

RESULT 12
AX105078 AX105078 4279 bp DNA linear PAT 30-APR-2001
LOCUS Sequence 22 from Patent WO0123540.
DEFINITION AX105078
ACCESSION AX105078
VERSION AX105078.1 GI:13921228
KEYWORDS
SOURCE dog:
ORGANISM Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE 1 (bases 1 to 4279)
AUTHORS Stocker, P.J., Steimel-Crespi, D.T., Crespi, C.L., Reif, T.C. and
Patten, C.J.
P-glycoproteins and uses thereof
Patent: WO 0123540-A 22 05-APR-2001;
GENTEST CORPORATION (US)
FEATURES
Location/Qualifiers
source 1..4279
/organism="Canis familiaris"
/db_xref="taxon:9615"
CDS
17..3862
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC37721.1"
/db_xref="GI:13921229"
/translation="MDPEGRKGSAEKNFMKMGKSKKNEKKEKKTSTVTFAMFRYSN
WLDRLYLVLGVTMAAIIHGAALPLMLLVFNMTDSFANAGISRNKTFPVIINISNTNT
OHFINHLEEMTVAAYYSYEGVAGVVAAYIOVSFWCLAGQAIILKIRKQFHAMKRLQ
ETGWFVDHVDGELNTRLTDDVSKINEGIDKIMGFQSIATPTFTKIVGTGKWLTL
VILASPVILGSAAIWAKILSPTDKELLAYAKAGAAVEVLAARTVIAFGQKKEL
ERYKNLEBEAGIGIKKAITANISGAAPFLIYALAFWTGTSILVSISGSGQVL
TVFFSLVGAFSIGQASPIEAFANARGAAVEYFIIDNKPSIDYSKSGHKNPKING
NLEFNKVHFSYPSRKEVILKGLNLKVSQGTVALVGNSSGCKSTVQLMQLYDPTD
GMVCTDGDQIRTNVHLREITGVVSQSPVLFAATIAENIRYGRNVMTDEKAVKE
ANAYDFIMKLPNKFTLNGERGAOLSGGKORIAARALVRNPKILLDEASALDTE
SEAVQVALDKARKORTTIVTAHRLSTVRNADVIAGDFDGVIVEKGNDELMEKGIY
FKLVTMQTRGNEIELENATGESSEDALEMSPKDSGSLIKRTRRSIHAPQODR
KLGTEDLENENVPVPSFWRLKLNSTWEPYFVVGIFCAIINGGLQPAFSIIISRIIGI
FTRDDEPTKQNSMFLVLGIIISFITPFLQGTFGKAGELITKELRLTVFRSM
LRQDYSWFDPPKNTTALTTLANDAAOVKAIGSLAVITONILNGTGIISIIY
WOLTLILLAIVIPILAIAGVEMKMLSGOALKDKKELEGAKTIAENFRVTVSLTR
EQKFYMYAQSLOVPYRNLKHAHIFGVFSFSTQMMYPSYAGCFRFGAYLVANFPMN
FQDVLVLFSAIVFGNAYVQVSSFPADYAKAVKSAAHVIMIIIEKPLDISYSPHGLKP
NTLEGNVTFNEVFNYPTRDIPVLQGLSLELVKGGTLLALVGSSGCGKSTVVQLLERF
```


YDPLAGSVLIDGKEIKHNLVQWRAHLGIVSQBPILFDCSTAEINAYGDNSRVVSGHEE
IMQAAKEANITHFTETLPEKYNTFVGDGTQSGGQKQRIATARALVRPHILLLDEA
LAKGDTSESKVQWQALDKAREGTCIVIAHRLSTQIADLIVVFQNGKVKREHGTQOOL
TSOQDTFVFSMVSVQAGAKR"

BASE COUNT	1295 a	833 c	1008 g	1143 t
ORIGIN				

Alignment Scores:		
Pred. No.:	5,21e-222	Length: 4279
Score:	354.00	Matches: 682
Percent Similarity:	74.74%	Conservative: 239
Best Local Similarity:	55.40%	Mismatches: 268
Query Match:	59.12%	Indels: 42
DB:	6	Gaps: 8

US-09-873-409-6 (1-1195) x AX105078 (1-4279)

Qy	1	MetIleuGlyIleLeuAlaSerLeuValAenGlyAlaCysLeuProLeuMetProLeu	20
Db	167	ATGTTGGTGGGCAATAGCTGCCATCATCCATGAGCTGCACCTCCTCATGATCGT	226
Qy	21	ValLeuGlyGluMetSerAspAenLeuIleSerGlyCysLeu-	34
Db	227	GTTTTGGAAACATGACAGATAGCTTGCMAATGAGAAATTCAGAAACAAACTTTT	286
Qy	35	-----ValGlnThrAsnThrTyrSerPhePheArg-	44
Db	287	CCAGTTATTAATTAATGAAGATATTACGAACAATACACAACATTTTCATCAACCATCTGGAG	346
Qy	45	-----LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe	59
Db	347	GAGGAAATGACACCATGATGCCTATTATTACAGTGGGATCGGTGCGCTCGTCTGGCTG	406
Qy	60	GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg	79
Db	407	GCTTACATCCAGGTTCATCTCTGGTCCCTGGCAGCAGGAACAGACATACTCAAAATTAGA	466
Qy	80	LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle	99
Db	467	AAACAATTTTTTCATGCTATCGCACAGGAGATTGGCTGGTTTGACGTGCATGACGTT	526
Qy	100	GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp	118
Db	527	GGGGAGCTTAACACCCGGCTCACAGACGATGCTCCAAATTCATTAAGAGGAATTTGGCGAC	586
Qy	119	LysIleAlaLeuLeuPheGlnAenMetSerThrPheSerIleGlyLeuAlaValGlyLeu	138
Db	587	AAAATGGATGTTCTTCAATCAATAGCAACATTTTCACCGGTTTTATATGCGGGTTT	646
Qy	139	ValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer	158
Db	647	ACACGTGGTTGGAAGCTAACCCCTTGATTTTGGCCATCAGCCCTGTTCTTGGACATTCA	706
Qy	159	AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer	178
Db	707	GCCGCCATCTGGGCAAGAATCATCTCTCATTTACTGATAAAGAACTTCGCGCTATGCA	766
Qy	179	LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg	198
Db	767	AAAGCTGGACGAGTAGCTGAAGAAGTCTTTAGCAGCAATCAGAACTGTGATTGCCCTTTGGA	826
Qy	199	AlaGlnGluLysGluIleuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly	218
Db	827	GGACAAAGAAGAAAGACTTGAAGAGGTACACAAAAATTTAGAAGAGACTTAAGGAATTTGGG	886
Qy	219	IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGly	238
Db	887	ATAAAGAAAGCTATACGGCCACAATTTCTATTGGTGGCGCTTCTTATTGATCTATGCA	946
Qy	239	ThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr	259
Db	947	TCATATGCTCTGGCTTCTGGTATGGGACTCTCTGGTCTCTCCAGTGA-----TAT	1000


```
Db 2075 AAGAGGACTTGAATGAGATGACCTCCAGTTCTCTGGAGGATTCTGAAGCTGAAC 2134
QY 624 LysProGluTyrProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 643
Db 2135 TCAACTGATGGCTTATTTTGGTGGTGGTATATTGCTATTATTAACGAGGACCTG 2194
QY 644 HisProValPheSerLeilePheAlaLysleileThrMetPhe--GlyAsnAsnAsp 662
Db 2195 CAACCCAGCATTTTCAATAATATTTTCAAGGATTATAGGGATTTTACCCGAGATGAGGAT 2254
QY 663 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 682
Db 2255 CCTGAACAAACACAGACAGATAGTATTAATGTTCTGTATTTCTAGTCCTTGGAAAT 2314
QY 683 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 702
Db 2315 ATTTCTTTTATACATTTTCTCCAGGGCTTCACATTTGGCAAGCTGGGAGATCCIC 2374
QY 703 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetIleuTyrGlnAspIleAlaTyrPhe 722
Db 2375 ACTAAGCGGCTTCGATACATGGTTTTCAGATCCATGCTGACAGCAGGATGTCAGCTGGTT 2434
QY 723 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 742
Db 2435 GATGACCCCTAAACACACCACTGGAGCATTCGACCAACAGGCTTCCTCAATGATGCGGCTCA 2494
QY 743 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 762
Db 2495 GTTAAAGGGCTATAGTTTCCAGCTGTGCTATACCCAGAAATATAGCAATCTTGGG 2554
QY 763 LeuSerValIleIleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIle 782
Db 2555 ACAGGCATTATATATCTTAATCTATGTTGGCAATTAACACTTTTACTCTTAGCAATT 2614
QY 783 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 802
Db 2615 GTACCCATCATTCGAATGACGAGGATGTTGAAATGAAATGTTGTGTGACCAAGCACTG 2674
QY 803 LysAspLysGlnLeuLysHisAlaGlyValIleAlaThrGluAlaLeuGluAsnIle 822
Db 2675 AAGATGAAGAAAGAGCTAGAGAGCTGGGAAGATGCTACAGAAGCCATCGAAAACTTC 2734
QY 823 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 842
Db 2735 CGAACTGTTCTTTCATCGGAGCAGAGTTTGAATACATGATGATGCACAGAGTTTG 2794
QY 843 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 862
Db 2795 CAAGTACCATACAGAACTCTTTGAGGAAAGCACACATCTTCGGGGTCTCATTTTCTATC 2854
QY 863 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 882
Db 2855 ACCGAGCAATGATGATTTTCTATGCTGGCTGTTTCCGGTGTGCTGCTACTTGGTG 2914
QY 883 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 902
Db 2915 GCAATGAGTTCATGAACCTTCAGGATGTTCTTTGGTATCTCAGCTATGCTTTGGT 2974
QY 903 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 922
Db 2975 GCCATGCGAGTGGGCGAGGTCAGTTTCATTTGCTCCTGACTATGCCAAAGCCAAAGTATCA 3034
QY 923 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 942
Db 3035 GCAGCCACGCTCATGATGATTCATTTGAAAGAGCCCTCTGATGACAGCTACAGCCCTCAC 3094
QY 943 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 962
Db 3095 GGCTCTAAGCCAAATACGTTGGAGAGGAAATGTGACATTTAATGAGGTCGTGTTCAACTAT 3154
QY 963 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 982
```

```
Db 3155 CCCACTCGACACAGACATCCCCGTGCTCCAGGGGCTGAGCCTCGAGTGAAGAGGCCAG 3214
QY 983 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 1002
Db 3215 ACGTGGCCCTCGTAGGTAGCAGTGGCTGTGGAGAGCACAGTTGTTTCAGCTCTCTAGAG 3274
QY 1003 ArgLeuTyrAspProValGlnGlnGlyValLeuPheAspGlyValAspAlaLysGluLeu 1022
Db 3275 CGCTTCTATGACCCCTGGCTGCTTTCAGTGAATTTGATGGCAAGAGATAAAGACACCTG 3334
QY 1023 AsnValGlnTyrLeuArgSerGlnIleAlaIleValProGlnGlnProValLeuPheAsn 1042
Db 3335 AATGTCAGTGGCTCGACACACCTGGGCATCGTGTCTCAGGAGCCCATCTCTGTTTGAC 3394
QY 1043 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 1062
Db 3395 TGCAGCATTTGCCGAGAACATTGCTTATGGAGACACACAGCCGGTCTGATACATGAAGAG 3454
QY 1063 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 1082
Db 3455 ATTATGACGGCAGCCCAAGGAGGCCAACATACACCACTTTCATCGAGACACTCCCTGAGAA 3514
QY 1083 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 1102
Db 3515 TACAACCCAGCAGTAGGACAGAAAGAACCCAGCTCTCTGTGGCCAGAAACAGCCCAT 3574
QY 1103 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 1122
Db 3575 GCCATAGCTCGGCTCTTCTGTAGACAGCTCATATTTTCTGTTGGATGAAGTACATCA 3634
QY 1123 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 1142
Db 3635 GCTCTGATACAGAAAGTGAAGAGTTGTCCAAGAGCCCTGGACAAAGCCAGAGAGGC 3694
QY 1143 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 1162
Db 3695 CGACCTGTGATGTGATCGCCACCGCTTGTCCACCATCCAGAAATGCAGATTTAATAGTG 3754
QY 1163 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 1182
Db 3755 GTGTTTCAGATGGCAAGTCAAGGACATGGCACACATCACAGCTGTGGGCCAGAA 3814
QY 1183 AspIleTyrPheLysLeuValAsnAlaGlnSer 1193
Db 3815 GGCATCTATTTTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3847

RESULT 13
AX105080 AX105080 4279 bp DNA linear PAT 30-APR-2001
LOCUS Sequence 24 from Patent WO0123540.
DEFINITION AX105080
ACCESSION AX105080
VERSION AX105080.1 GI:13921230
KEYWORDS dog.
SOURCE Canis familiaris
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
AUTHORS 1 (bases 1 to 4279)
Stockert,P.J., Steimel-Crespi,D.T., Crespi,C.L., Reif,T.C. and
Patten,C.J.
TITLE P-glycoproteins and uses thereof
JOURNAL Patent: WO 0123540-A 24 05-APR-2001;
GENTEST CORPORATION (US)
FEATURES
source Location/Qualifiers
1..4279
/organism="Canis familiaris"
/db_xref="taxon:9615"
17..3862
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC37722.1"
/db_xref="GI:13921231"
/translation="MDPEGGRKGSAAEKFNFKMGKSKKKEKKEKPTVSTFAMFRYSN"
```


Qy	539	AlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuVal	558
Db	1841	GTGGAGAAAGGAAATCATGATGAATCATGAAAGAGAGAGGCGATTTACTTCAAACTGTGC	1900
Qy	559	MetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu	578
Db	1901	ACAATGCAG---ACAAGGAAATGAAATGTAGTTAGAAAATGCCACTGGTGAATCCAAA	1957
Qy	579	ArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle---	597
Db	1958	AGTGAAAGTGCATCCCTTGGAAATG---TCTCCAAAAGATTCCAGGCTCCAGTTTAATAA	2014
Qy	598	-----ApLysAlaGluLysSer	603
Db	2015	AGAAATCAACTCCGAGGAGTATACATGCACCAAGGCCAAGACAGAGAACTTGGTACA	2074
Qy	604	ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn	623
Db	2075	AAAGGAGCTTGAATGAGAAATGACCTCCAGTTTCTCTCGAGAGATTCTGAGCTGAC	2134
Qy	624	LysProGluTyrProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal	643
Db	2135	TCAACTGAATGGCTTATTTTGTGGTGTATATTTTGTGCTATTATAAACGGAGGCTG	2194
Qy	644	HisProValPheSerIleIlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAsp	662
Db	2195	CAACGAGCATTTTCAATAATATATTTTCAAGGATATATAGGATCTTTTACCACGAGATGAGGAT	2254
Qy	663	LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal	682
Db	2255	CCTGAAACAAACGACAGAAATGATACATGTTTTCTGTATTGTTTCTTGATCTCTTGAA	2314
Qy	683	IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu	702
Db	2315	ATTTCTTTTATTACATTTTCTCCAGGCTTTCACATTTGGCAAGCTGGGAGATCCTC	2374
Qy	703	ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPhe	722
Db	2375	ACTAAGCGGCTTCGATACATGTTTTTCAGATCCATGCTGAGACAGGATGTCCAGCTG	2434
Qy	723	AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln	742
Db	2435	GATGACCTTAAACACCACTGAGGATTTGACACACAGGCTTGCCATGATCGGCTCAA	2494
Qy	743	IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly	762
Db	2495	GTTAAGGGGCTATAGTTCCAGCTTGCTGTCATTTACCAGAAATATAGCAAACTTTGGG	2554
Qy	763	LeuSerValIleIleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIle	782
Db	2555	ACAGGCAATATATATCCTTAAATCTATGTTGGCAATTAACACTTTTACTCTTAGCA	2614
Qy	783	AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn	802
Db	2615	GTACCCATCATTCGAATAGCAGAGTTGTGAAATCGAAAATGTTCTGTGGACAGCACTG	2674
Qy	803	LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle	822
Db	2675	AAAGATAAGAAAGAGCTAGAAAGAGAGCTGGGAAGATTGCTTACAGAGCCATCGAAA	2734
Qy	823	ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu	842
Db	2735	CGAATGTTGTTCTTTTGATCTCGGAGCAGAGTTTGAATACATGATGTACACAGATTG	2794
Qy	843	GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe	862
Db	2795	CAAGTACCATACAGAAACTCTTTGAGGAAGCACACATCTTCCGGGCTCTCATTTTCTATC	2854
Qy	863	SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle	882
Db	2855	ACCCAGGCAATGATGATATTTTCTATGCTGGCTGTTTCCGTTTGGTGCTACTTGCTG	2914

Qy	883	GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyGly	902
Db	2915	GCAATGAGTTCACTAACTTTCCAGATGTTCTTTGGTATTCTCTCAGCTATTGTCTTTGGT	2974
Qy	903	AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyTyrSerLysAlaLysSerGly	922
Db	2975	GCCATGGCAGTGGCGCAGTCACTTTCATTTGCTCTCTGACTATGCCAAGACCCAAAGATCA	3034
Qy	923	AlaAlaHisLeuPheAlaLeuLeuGluLysTyrProAsnIleAspSerArgSerGlnGln	942
Db	3035	GCAGCCACGTCATCATGATTCATTGAAAAGGCCCTCTGATTGACAGCTACAGCCCTCAC	3094
Qy	943	GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr	962
Db	3095	GGCCTCAAGCCAAATCGTTTGGAGGAAATGTGACATTTAATAGGTCGTGTTCAACTAT	3154
Qy	963	ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys	982
Db	3155	CCCACTCGACAGACATCCCGTCTCCAGGGGCTGAGCCTCGAGTGGAGAGGGCCAG	3214
Qy	983	ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuGln	1002
Db	3215	ACGTGGCCCTCTGAGTAGCAGTGGCTGTGGAGAGCACAGTTGTTGAGCTCTCTAGAG	3274
Qy	1003	ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu	1022
Db	3275	CGCTTCTATGACCCCTGGCTGGTTCAGTGCCTAAATTGATGGCAAGATTAAGCACCTG	3334
Qy	1023	AsnValGlnThrLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn	1042
Db	3335	AATGTCAGTGGCTCCGAGCACACCTGGGCGATCGTGTCTCAGAGCCCATCTGTTTGAC	3394
Qy	1043	CysSerIleAlaGluAsnIleAlaTyTyrGlyAspAsnSerArgValValProLeuAspGlu	1062
Db	3395	TGCAGCATTCGCCGAGAACATTGGCTATGGAGACACACCGCGTCTCATCATGAAGAG	3454
Qy	1063	IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys	1082
Db	3455	ATTATGCGAGCAGCCAAAGGAGGCCAACATACACCACTTCATCTCGACACACTCCCTGAGAAA	3514
Qy	1083	TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu	1102
Db	3515	TACAACACCAAGTAGGAGACAAAGGAAACCCAGCTCTCTGTGGGCCAGAAACAGCGCAAT	3574
Qy	1103	AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer	1122
Db	3575	GCCATAGCTCGCGCTCTGTGTAGACAGCTCATATTTGCTTTTGATGAAGCTACATCA	3634
Qy	1123	AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly	1142
Db	3635	GCTCTGGATACAGAAAGTGAAGAGTTGTCCAAGAGCCCTGGCAAAAGCCAGAGAAAGGC	3694
Qy	1143	ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal	1162
Db	3695	GGCACCTGCATTTGTATGCGCCCAACCGCTGTGCCACATCCAGAAATGCAGATTTAATAGTG	3754
Qy	1163	ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlnLeuLeuArgAsnArg	1182
Db	3755	GTGTTTCAGAAATGCCAAGATCAGAGGAGCATGCCACACATCAACAGCTGCTGGCCAGAAA	3814
Qy	1183	AspIleTyrPheLysLeuValAsnAlaGlnSer	1193
Db	3815	GGCATCTATTTTCCATGGTCACTGTCAGTGTCCAGGCT	3847

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 Ruth, A. and Roninson, I.

Mutations of the mdrl p-glycoprotein that improve its ability to

confer resistance to chemotherapeutic drugs

Patent: WO 0210205-A 1 07-FEB-2002;

THE BOARD OF TRUSTEES OF THE UNIVERSITY OF ILLINOIS (US)

FEATURES

Location/Qualifiers

1..4646

/organism="Homo sapiens"

/db_xref="taxon:9606"

425..4267

/note="unnamed protein product"

/codon_start=1

/protein_id="CAD28599.1"

/db_xref="GI:19584229"

/translation="MDLEGRNGGAKKKNFFKLNKSEKDKKEKPTSVFVMFRYSN

WLDKLVYVGTLLAAIIFHAGLPLMLVFGEMTDIFANAGNLDELMSNITNRSDINDTG

FMNLDEEDTRVAYVYGGIAGVLVAAYIQVSWFCLAGROJHKQFFHAIMRQEI

GHFDVHDGELATRLDDYSKINEGIGDKIGWFOFOSMATFFTGIVGTRGKWLTLVI

LALSPVIGSAUAWKILASPTDKELLAYAKAGAVAEVLAIRTVIAPGGOKSELR

YKNLEAKRIGIKAITANSISGAAPFLIYASALAFWYGTLLVLSGEYSIQGLTV

FFSLTIGAFSVGQASPIEFANARGAAEYIFKIDNKPSIDSYSGHGPNDIKGNL

EFNRHFSYPSRKEVILKGLNKLKVGQGTVALVNGSGCKSTTQMLQRLYDPTGEM

VSDGQDITRTINVRFLREIIGVVSQEPVLPAITIAENIRYGRNVTMDIEKAVKAN

AYDPIMKLPHKEDTLVGERGAOLSGGOKORIAALAVNPVKILLDELALDTESE

AVVQVALDKARKGRTIIVIAHRLSTVRNADVAGDDGVIVKGNHDELMKEGIYFK

LVTMQAGNELENADESKSEIDALEMSNDRSLIRKRSKTRSVRSGQARQKRL

STKEALDESIPVPSFRIMKLNLTWPYFVWGVFCALINGIQLQAPALIFSKIIVFT

RIDDPTRQNSNLSFLFLALGIIISFTFLQGTFFGKAGELTKRLRYVMVFRMLR

QDVSWFDKPNKTLRLANDAAQVKAIGSRLAVITONIANLGTGIIISFIYGMQ

LTLLALLIPIIAIAGVEMKLSGOALKDKKEGAKIATEAENRPTVLSLTQEO

KEHMYAQSLQPYRNSLKAHIFGTFSTQAMYPFSVAGCFERGVIAVAKLSFE

DVLLVFSAVFVGMAVQVSSFAPDIYAKAKISAHLIMIEIKTPLIDSYSTGLMPT

LEGNVTFFGVFNYPTRPIVLQGLSEVQKPTLALVSGSGCGKSTVQLLERFYD

PLAKVTLGDKETIKRLNVOWLRAHLGIVSQEPIFLFDCSAENIAGDNSRVVSOEIV

RAAKENIHAFTSPKYSTKVGDKGTOLSGGOKORIAALAVNPVKILLDELALD

ALDTESEKVVQEALDKARSGRTCIIVIAHRLSTIYNADLIVVQNGRVEHGHQQLA

QKGIYSMSVQAGTKRQ"

BASE COUNT 1372 a 892 c 1130 g 1252 t

ORIGIN

Alignment Scores:

Pred. No.: 5 78e-222 Length: 4646
Score: 3554.00 Matches: 682
Percent Similarity: 75.63% Conservative: 246
Best Local Similarity: 55.58% Mismatches: 263
Query Match: 59.12% Indels: 36
DB: 6 Gaps: 9

US-09-873-409-6 (1-1195) x AX391099 (1-4646)

Qy 1 MetileLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20

Db 575 ATGGTGGGGAATTTGGCTGCCATCATCCATGGGCTGGACTTCTCTCATGATCGT 634

Qy 21 ValLeuGlyGluMetSerApp-----AsnLeuIleSer 31

Db 635 GTGTTTGGAGAAATGACAGATATCTTTGCRAAATGACAGAAATTTAGAGATCTGATGTC 694

Qy 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43

Db 695 AACATCTAATAGAGATGATCAATGATACAGGGTCTTCATGAATCTGGAGGAAGAC 754

Qy 44 -----ArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61

Db 755 ATGACAGGATGCTATTATTACAGTGAATTTGGTCTGGGGTGGTGGTCTGCTTAC 814

Qy 62 IleGlnIleSerLeuTyrIleleThAlaAlaArgGlnThrLysArgIleArgLysGln 81

Db 815 ATTCAGGTTTCATTTTGGTGGCTGGCAGCTGGAAGACAAATACACAAAATTTAGAAAA 874

Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101
Db 875 TTTTTCATGCTATAATGCGACAGAGATAGCTGTTTGTATGTCACGATGTTGGGAG 934

Qy 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db 935 CTTAACACCCGACTTACAGATGATCTCTAAGATTAATGAAGGAATTTGGTGAACAAAT 994

Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db 995 GGAAATGTTCTTCAGTCAATGCAACATTTTTCACATGGGTTTATAGTAGATTTACAGCT 1054

Qy 141 GlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaIle 160
Db 1055 GGTGGGAAGCTAACCCCTTGATTTTGGCCATCAGTCTCTGTTCTTGGAGCTGTGAGCTGCT 1114

Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db 1115 GTCTGGGCAAGATATACTATCTTCTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 1174

Qy 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db 1175 GGAGCAGTAGCTGAAGAGTCTTGGCAGCAATTAGAAGTGTGATTCATTTGGAGACAA 1234

Qy 201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
Db 1235 AAGAAAGAACTTTGAAAGGTACAAACAAATTTTGAAGAAGCTTAAAGAAATTTGGGATAAAG 1294

Qy 221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
Db 1295 AAGACTTATTACAGCCAAATTTCTTATAGTGTGCTTCTCTGCTGATCATCTTAT 1354

Qy 241 GlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
Db 1355 GCTTGGCTTCTGATGAGGACCACTTGGTCTCTCAGGGAA-----TATTCATT 1408

Qy 261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
Db 1409 GGACAAAGTACTCACTGATTTCTTTCTGATTAAATTTGGGGCTTTTATGTTGTGACAGGCA 1468

Qy 281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
Db 1469 TCTCAAGCATTGAAGCATTTCGAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1528

Qy 301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
Db 1529 ATTGATAAATAGCAAGTATTGACAGCTATTGCAAGAGTGGGACAAACACAGATAATTT 1588

Qy 321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
Db 1589 AAGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAAGTTAAG 1648

Qy 341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
Db 1649 ATCTGAAGGGCTCAACCTGAAGGTGCGAGTGGCAGACGGTGGCTGGTGGTGGAAAC 1708

Qy 361 AsnGlySerGlyLysSerThrValValGlnLeuGlnArgLeuTyrAspProAspAsp 380
Db 1709 AGTGGCTGTGGGAAGAGCACACAGTCCAGCTGATGAGAGGCTCTATGACCCACAGAG 1768

Qy 381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
Db 1769 GGGATGTCAGTCTTCATGGACAGGATATTAGGACCATTAATGTAAGGTTTCTACGGAA 1828

Qy 401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
Db 1829 ATCAATGGTGTGGTGTGAGTACAGAACCTGTATTGTTTGGCCACCACTAGCTGTAACAT 1888

Qy 421 LysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
Db 1889 CGTATGGCCGTGAAATGTCCATGATGATGATGAGAAAGCTGTCAAGGAGGCAAT 1948

Qy 441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyLysGly 460

Db	1949	GGCTTATGACCTTTATCATGAAGAACTGCCTCATTAATAATTGCACCCTCGTGTGGAGAGAAGGG	2008
Qy	461	AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn	480
Db	2009	GCCAGTGTGAGTGGTGCCGAGACAGAGATGCCCATTCGACGTGCCCTGGTTCGCAAC	2068
Qy	481	ProLysIleLeuIleLeuAepGIuaIathrSerAlaLeuAapSerGluSerLysSerAla	500
Db	2069	CCCCAAGATCCTCGCTGGTATGAGGCCAGCTCAGCTTGGCACAGAAAAGCGCAAGACAGT	2128
Qy	501	ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValAlaIleHisArg	520
Db	2129	GTTCCAGTGGCTCTCGATAGGCCAGAAAGGTGCGACCATTCGTGATAGCTCATCGT	2188
Qy	521	LeuSerThrIleArgSerAlaAepLeuIleValThrLeuLysAepGlyMetLeuAlaGlu	540
Db	2189	TTGTCACAGTTCGTAATGCTGACGTCATCGCTGGTTCGATGATGGAGTCATTGTGCGAG	2248
Qy	541	LysGlyAlaHisAlaGluLeuMetalLysArgGlyLeuTyTrSerLeuValMetSer	560
Db	2249	AAAGGAAATCATGATGAATCATGAAAGAGAACGCAATTTACTTCAAACATTTGTCCAATG	2308
Qy	561	Gln-----AspIleLysLysAlaAepGluGlnMetGluSerMetThr	574
Db	2309	CAGACAGCAAGAAATGAAGTGAATTAGAAAAATGCAAGCTGATGAATCCAAAAGTGAATT	2368
Qy	575	TyrSerThrGluArgLysThrAenSerLeuProLeuHisSerVal-----	589
Db	2369	GATGCCCTGGGAATGCTTCCMAATGATTCAGAGTCAGCTCTAATAAGAAAAAGATCAACT	2428
Qy	590	---LysSerIleLys---SerAaspPheIleAaspLysAlaGluSerThrGlnSerLys	607
Db	2429	CGTAGGAGTGTCCGTGGATCACAAGCCCAGACAGAAAGCTTAGTAGTACCAGAGGCTCTG	2488
Qy	608	GluIleSerLeuProGluValSerLeuLysIleLeuLysLeuAsnLysProGluTrp	627
Db	2489	GATCAAAGTAGTATACCTCCAGTTTCCTTTTCGGAGGATATAGAGCTAAATTTAACTCAAATG	2548
Qy	628	ProPheValValLeuGlyThrLeuAlaSerValLeuAenGlyThrValHisProValPhe	647
Db	2549	CCTTATTTGTTGTGTGTATTTTTGTGGCATATATAATGGAGGCTCGCAACACAGCATTT	2608
Qy	648	SerIleIlePheAlaLysIleIleThrMetPheGlyAsn--AsnAaspLysThrThrLeu	666
Db	2609	GCAATAATATTTCAAAGATTATAGGGTTTTTACAAGAATTTGATGATCCTGAAACAAA	2668
Qy	667	LysHisAaspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCySPheVal	686
Db	2669	CGACAGAAATAGTAACTTGTTTTCACATATTGTTCTAGCCCTTGGAAATTAATTTCTTTAT	2728
Qy	687	SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThMetArgLeu	706
Db	2729	ACATTTTTCCTTCAGGGTTTCCATTTGGCAAGCTGGAGAGATCCTCACCAGCGGCTC	2788
Qy	707	ArgHisLeuAlaPheLysAlaMetLeuTyTrGinAspIleAlaTrpPheAaspGluLysGlu	726
Db	2789	CGATACATGGTTTCCATCCATGCTTCAGACAGGATGTGAGTTGGTTTGATGACCCATAA	2848
Qy	727	AenSerThrGlyGlyLeuThrThrIleLeuAlaIleAspileAlaGlnIleGlnGlyAla	746
Db	2849	AACACCACCTGGAGCATTGACTCACCGCTCGCCAAATGATGTGCTCAAGTTAAAGGGCT	2908
Qy	747	ThrGlySerArgIleGlyValLeuThrGlnAenAlaThrAsnMetGlyLeuSerValIle	766
Db	2909	ATAGGTTTCCAGGCTTGCTGTAATTTACCCAGAAATATAGCAAAATCTTTGGGACAGGAAT	2968
Qy	767	IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu	786
Db	2969	ATATCTTTCATCTATGGTTGGCAACTAAACACTGTTACTCTTAGCAATTTGTACCATCAT	3028
Qy	787	AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln	806

D	b	3029	GCAATAGCAGGAGTTGTGGAAATGAATAAATGTTGCTCTGGACAAGACACTGAAAGATAGAATAAGAAA	3088
Q	y	807	GlulLeuLysHisAlaGlyIleLysIleAlaThrGluLaLeuLeuGluAsnIleArgThrIleVal	826
D	b	3089	GAACCTAGAAAGTGCTGGAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTCGTT	3148
Q	y	827	SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis	846
D	b	3149	TCTTTTGACTCAGGAGCAGAAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCACATAC	3208
Q	y	847	ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe	866
D	b	3209	AGAAACTCTTTGAGGAAGACACACTCTTTGGATTACATTTTCTTACCAGCGCATG	3268
Q	y	867	IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg	886
D	b	3269	ATGTATTTTTCTATGCTGATGTTTCCGGTTTGGAGCCTACTTGTGGCGCACATAAACTC	3328
Q	y	887	MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle	906
D	b	3329	ATGAGCTTTGAGGATGTTCTGTAGTAGTATTTTTCAGCTGTTGCTTTTGGTGCCATGCCGCTG	3388
Q	y	907	GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu	926
D	b	3389	GGGCAAGTCAGTTTCATTTGCTCTCTGACTATGCCCCAAAGCCAAATATCATGACGCCACATC	3448
Q	y	927	PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysPro	946
D	b	3449	ATCATGATCATTGNAAAAACCCCTTTGATTGACAGCATACAGCACGGAGGCTTAATGCG	3508
Q	y	947	AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro	966
D	b	3509	AACACAVTTGGAAGGAATGTACATTTGGTGAAGTTGTATTCAAATATCCCACCCGACCG	3568
Q	y	967	AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe	986
D	b	3569	GACATCCCAGTGCTTCAGGACATGAGCTGGAGGTGAAGAAGGGCCAGACGCTGCTG	3628
Q	y	987	ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuGlnArgLeuTyrAsp	1006
D	b	3629	GTGGCAGCAGTGGCTGTGGAGAGACACAGTGTCTCAGCTCTCTGGAGCGGTTCTACGAC	3688
Q	y	1007	ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp	1026
D	b	3689	CCCTTGGCAGGGAAGTGCTGCTTGATGGCAAGAAATAAAGCGACTGAATGTTCACTGG	3748
Q	y	1027	LeuArgSerGlnIleAlaIleValProGlnGluProValIleUpheAsnCysSerIleAla	1046
D	b	3749	CTCCGACGACCTTGGGCATCTGTGCTCCAGGAGCCCATCTCTTTGACTGCAGCATTGCT	3808
Q	y	1047	GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluLysGluAla	1066
D	b	3809	GAGAACATTGGCTATGAGACACACAGCCGGTGGTGTCTACAGGAGAGAGNTCGTGAGGCA	3868
Q	y	1067	AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln	1086
D	b	3869	GCAAGGAGGCCAACATACATGCTTTCATCGAGTCACTGCCTTAATAAATATAGCACTAAA	3928
Q	y	1087	ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg	1106
D	b	3929	GTAGGAGACAAAGGAACCTACGCTCTCTGGTGGCCAGAAACAACGATGTGCCATAGCTCGT	3988
Q	y	1107	AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn	1126
D	b	3989	GCCCTTGTATGACAGCCCTCATATTTTGTCTTTGGATGAAGCCACAGTCAGCTCTGGATACA	4048
Q	y	1127	AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu	1146
D	b	4049	GAAAGTGAAGAGTTGTCCAAGAGCCCTGGACAAGCCAGAGAGAGGCCGCGCACCTGCAAT	4108
Q	y	1147	ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisLeu	1166
D	b	4109	GTGATGTCTACCCCGCTGTCCACCATCCAGAATGSCAGACTTAATAGTGTGTTCAGAAAT	4168


```
QY 1167 GlycylLeuValGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1186
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 4169 GGCAGAGTCAGAGGAGTCGACGATCAGCAGCTGCTGCGACAGAAAGCACTATTTT 4228

QY 1187 LysLeuValAsnAlaGlnSer 1193
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 4229 TCAATGTCAGTGCAGGCT 4249

RESULT 15
AF269224 4045 bp DNA linear SYN 04-JUL-2000
LOCUS Synthetic construct Canis familiaris his-tagged-multidrug
DEFINITION resistance glycoprotein gene, complete cds.
ACCESSION AF269224
VERSION AF269224.1 GI:8926216
KEYWORDS synthetic construct.
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1 (bases 1 to 4045)
AUTHORS Roulet,A.C., Puel,O., Gesta,S., Alvinerie,M. and Pineau,T.J.
TITLE Analysis of dog MDR1 p-glycoprotein
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 4045)
AUTHORS Roulet,A.C., Puel,O., Gesta,S., Alvinerie,M. and Pineau,T.J.
TITLE Direct Submission
JOURNAL Submitted (19-MAY-2000) Pharmacology, INRA, 180 Chemin de
Tournepieuille, BP3, Toulouse 31931, France
FEATURES
source
/organism="synthetic construct"
/db_xref="taxon:32630"
/notes="original mRNA isolated from Beagle dog colon"
101..3964
/codon_start=1
/transl_table=1
/product="his-tagged-multidrug resistance glycoprotein
MDR1"
/protein_id="AAF81747.1"
/db_xref="GI:8926217"
/translation="MHFHHHPDGGKRGSAEKNFWMGKSKKNEKKEKPTVSTFA
MFRYSNTHLNLVMTMAAIIHGAAIPLMVLVFGNMTDSFANAGISRNKTPFVINE
SITNTQHTINLEEMTVAAYVYSGIGAGVLVAAVYQVSWCLAAGRQILKIRKOFF
HAIMROEIGFVDHVDGEINTRLTDDVSKINEGIDGKGMFFQSIATFTFTVGFTR
GWLKTLVLAISPLVLSAIAWAKILSSSTDRELLAYAKAGAAVEVLAIETVIAFG
GQKELEYNKNLEERKIGIKAITANISIGRAFLIYASIALAFWYGTSLVLSSEY
SIGQVLTVFVSLGAFSGQASPIEAFANARGAAVEIFPKIDNPKSIDSYSKSHK
PNINKLEFFKNVHFSPYRKEVKILKLNKVSQOTVALVNGSCGSTTVQLMQR
LYDPTDGMVCIQDQIRTNVRLHREITGVVSOEPLVFATTAENIRYGRNVWDEI
EKAVKENAVDFTMKLPNKFDTLVGERGAQLSGGOKORTAIARALVRNPKILLDEAT
SALDTSEAVQVALDKARKRTTIVIAHRLSTVRNADVIAGFDGVIIVEKGNHDELM
KEGIVFKLVMTQTRNEIELENTGESKESDALEMSPKDSGSLIKRSTRSLIHA
POQDRKLGKTDLENENPVVPSFWRILKLNSTENVPYVVGIFCAINGLQPAFSIF
SRIGITFDEDEDPETRKSNMPSVLFLVLIIGISIFTFLOGTFKGAEIILTKRLY
MVRPSLRDQVDFDQNTGALTTRLANDAOKVGAIGSLAVITONIANLGTGII
ISLIGWQTLILAIIVPIIAIAGVEMKMSQALKDKKELEGAKIATELENPT
VNSLTREQPEYMAQSLQVPRNSLRKAHIFGVSFSTIQAMMYFYACFPQFAYLV
ANEFMFPDVLVFSIAIVFGAMAVQVSPADYAKAVSAAHVIMIEKSPILDSYS
PHGLKENTGNTLVNEVFNFTREDIPVLOGLSLEVKKQTLALVSGSGCKSTV
QLLERVYDPLAGSLVDGKEIKHLNWLRLHGVISQEPILFDCAENIAYGNRSR
VVSHERIMAOEKANITHFTETLPEKYNTRVGGKGTQLSGGOKORTAIARALVROPHI
LILDEATSALDTSEKVVQEQALDKAREGRTCVIAHRLSTIQNADLIIVVFQNGKVEH
GTHQQLLAQKGIYFSMISVQAGKR"
BASE COUNT 1171 a 850 c 986 g 1038 t
ORIGIN

Alignment Scores:
Pred. No.: 6 54e-222 Length: 4045
Score: 3552.00 Matches: 681
Percent Similarity: 74.74% Conservative: 239
Best Local Similarity: 55.32% Mismatches: 269
Query Match: 59.08% Indels: 42
```


Db	1283	AAATATTAAAGGGAATTTGGAAATTCAAAAATGTTTCACTTCAGTTACCCCTTCGAAAAAGAA	1342
Qy	339	IleYsIleLeuLysGlyLeuAenLeuAArgIleLysSerGlyGluThrValAlaLeuVal	358
Db	1343	GTTAAGATCTTAAAGGTCACACCTGAAGTTTCAGATGGGCAGACAGTGGCGCTGGT	1402
Qy	359	GlyLeuAenGlySerGlyLysSerThrValValGlnLeuLeuGlnAArgLeuTyAspPro	378
Db	1403	GGGAACAGTCGGCTGGCGGAAGACACACCGTCGAGCTGATGCAGAGGCTCATGACCCC	1462
Qy	379	AspAspGlyPheIleMetValAspGluAenAspIleAArgAlaLeuAenValAArgHisTy	398
Db	1463	ACAGATGGCATGTCTGTATTGTATGGACAGGACATTTAGGACCATTAATATTGAAGGCATCTT	1522
Qy	399	ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn	418
Db	1523	CGGGAATTAATCTGGTGTGGTGAGTCAGGAGCCTGTGTGTTGCCACACAGATAGTGAA	1582
Qy	419	AsnIleLysTyGlyArgAspAspValThrAspGluGluMetGluAArgAlaAArgGlu	438
Db	1583	AACATTCGCTATGCCCGCAAAATGTCACATGGATGAGATTGAGAAAGCTGTTAAAGAA	1642
Qy	439	AlaAsnAlaTyAspPheIleMetGluPheProAenLysPheAsnThrLeuValGlyGlu	458
Db	1643	GCCAATGCCCTATGATTTATCATGAAACATACCCTAAATAATTTGACACCTCTGGTTGGAGAG	1702
Qy	459	LysGlyAlaGlnMetSerGlyGlyGlnLysGlnAArgIleAlaIleAlaAArgAlaLeuVal	478
Db	1703	AGAGGGCCACAGCTGAGTGGTGACAGAAACAGAGNATCGCCATTCCTCGGGCCCTGGT	1762
Qy	479	ArgAsnProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys	498
Db	1763	CGCAACCCCAAGATCTCTCTCGTGGATGAGGCAACGTCAGCTCTGCACACTGAAAGTAA	1822
Qy	499	SerAlaValGlnAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAla	518
Db	1823	GCAGTGGTTTCAGGTGGCCCTGGATAAGGCGCAGAAAGGCGGACTACCATTTGTCATAGCT	1882
Qy	519	HisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeu	538
Db	1893	CATCGTTGTCTCAGTTCGTAATGCCGATGTCATTCGTGGTTTGTATGGATGCATTT	1942
Qy	539	AlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyTySerLeuVal	558
Db	1943	GTGGAGAAAGAAATCATGATGAATCATGAAAGAGAGAGGCACTTACTTCAAACTTTGTC	2002
Qy	559	MetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTySerThrGlu	578
Db	2003	ACAATGTCAG--ACAAGAGAAATGAAATTTGAAATTTAGAAATTCGCACTGGTGAATCCAAA	2059
Qy	579	ArgLysThrAenSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle---	597
Db	2060	AGTGAAGTGTATGCTTGGAAATG---TCTCCAAAGATTCAGGTCACGTTTAATAAAA	2116
Qy	598	-----AspLysAlaGluCLeuSer	603
Db	2117	AGAAGATCAACTCGCAGGAGTATACATGCACACCAAGGCGCAAGAGCTTGGTGACA	2176
Qy	604	ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAen	623
Db	2177	AAAGAGACACTTGAATGAGAATGTACCTTCAGCTTCCAGTTTCTCTCGAGGATTCGAAAGCTGAAC	2236
Qy	624	LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAenGlyThrVal	643
Db	2237	TCAACTGAATGGCCCTATTTTGGTGGTATATTTTGGTCTATTATTAACCGAGGCCCTG	2296
Qy	644	HisProValPheSerIleIlePheAlaLysIleIleThrMetPhe--GlyAsnAsnAsp	662
Db	2297	CAACAGCATTTTCAATAATATTTTCAAGNATTAAGGATCTTTACCCGAGATGAGCAT	2356
Qy	663	LysThrLeuTyHisAspAlaGluIleTySerMetIlePheValIleLeuGlyVal	682

2357	Db	CCTGAAACAAAACGACAGAAATAGTAACATGTTTCTCGTATTGTTTCTAGTCCTTGAAATT	2411
683	Qy	IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu	702
2417	Db	ATTTCTTTTATTACATTTTTCTCCAGGGCTTCACATTTTGGCAAGCTGGGAGATCCTC	2476
703	Qy	ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPhe	722
2477	Db	ACTAAGCGCTTCGATACATATGTTTTCAGATCCATGCTCAGACAGAGATGTCAGCTGGTTT	2536
723	Qy	AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln	742
2537	Db	GATGACCTTAAACACCACTGGAGCATTTGACACACAGCGTTGCCAATGATCGGCTCAA	2596
743	Qy	IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly	762
2597	Db	GTTAAAGGGGCTATAGTTCCAGGCTTGCTCATTTACCCAGAATATAGCAAAATCTTGGG	2656
763	Qy	LeuSerValIleIleSerPheIleTyrGlyTyrGluMetThrPheIleuIleLeuSerIle	782
2657	Db	ACAGGCATTTATTATATCTTAATCTATGTTGGCAATTAACACTTTTACTCTTAGCAATT	2716
783	Qy	AlaProValIleAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn	802
2717	Db	GTACCCCATCTTGCAATACAGGAGTTGTTGAAATGAAATGTTGTCTGGCAAGCACTG	2776
803	Qy	LysAspLysGlnIleuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle	822
2777	Db	AAAGATAAGAAAGAGCTAGAAAGGAGCTGGAGAGTTGCTACAGAAGCCATCGAAAACTTC	2836
823	Qy	ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu	842
2837	Db	CGAACTGTGTTCTTTTGACTCGGAGACAGAGTTTGAATACATGATGTCACAGAGTTTG	2896
843	Qy	GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe	862
2897	Db	CAAGTACCATCAGAAAACTCTTTGAGGAAGACACATCTTCGGGTCTCATTTTCTATC	2956
863	Qy	SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle	882
2957	Db	ACCCAGGCAATGATGATATTTTCCCTATGCTGGCTGTTTCCGGTTTGGTGCTACTTGGTG	3016
883	Qy	GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly	902
3017	Db	GCAAAATGAGTTTCATGAACCTTTCAGGATGTTCTTTTGGTATTTCTCAGCTATGTCTTGGT	3076
903	Qy	AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly	922
3077	Db	GCCATGGCAGTGGGCGAGTCAGTTCATTTGCTCTCGACTATGCCAAGCAAGATATCA	3136
923	Qy	AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu	942
3137	Db	GCAGCCACGTCATCATGATCATTGAAAAAGGCCCTCTGTATTCACAGCTACAGCCCTCAC	3196
943	Qy	GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr	962
3197	Db	GGCCTCAAGCCAAATACGTTGGGAAGAAATGTGACATTTTAAATGAGGTCGTGTTCACATAT	3256
963	Qy	ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys	982
3257	Db	CCCACTCGACACATCCCCGGTCTCGAGGGCTGAGCCTCGAGGTGAAGAGGGCCAC	3316
983	Qy	ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln	1002
3317	Db	ACGCTGGCCCTCGTAGGTAGCAGTGGCTGTGGGAAGACACAGTTGTTACGCTCTTAGAG	3376
1003	Qy	ArgLeuTyrAspProValGlnGlnValLeuPheAspGlyValaspAlaLysGluLeu	1022
3377	Db	CGCTTCTATGACCCCTTGGCTGTTGCTAATTTGATGGCAAGAGATAAAGCACCTG	3436
1023	Qy	AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn	1042
3437	Db	AATGTCAGTGGCTCCGACACACCTCGGCATCGTGTCTCAGAGGCCATCTCTGTTTGG	3496


```
QY 1043 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValProLeuAspGlu 1062
Db 3497 TGCAGCATTCGCCGAGACATTCCTATGGAGACAACAGCCGGTCTGTATCACATGAAGAG 3556
QY 1063 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 1082
Db 3557 ATTATGCAGGCAGCAGGAGGCCAACATACACCCTTCATCGAGACACTCCCTGAGAAA 3616
QY 1083 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 1102
Db 3617 TACAACACCCAGAGTAGGAGACAAAGGAACCCAGCTCTCTGGTGGCCAGAAACAGCCGCAAT 3676
QY 1103 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuAspGluAlaThrSer 1122
Db 3677 GCCATAGCTCGCGCTCTTGTAGACAGCCTCATATTTTGGTGGATGAAGCTACATCA 3736
QY 1123 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 1142
Db 3737 GCTCTGGATACAGAAAGTGAAGAGTTGTCCAAGAGCCCTGGACAAAGCCAGAGAGGC 3796
QY 1143 ArgThrCysLeuValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 1162
Db 3797 CGCACCTGCATTGTATCGCCACCGCTTGTCCACCATCCAGAAATGCAGATTTAATAGTG 3856
QY 1163 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlnLeuLeuArgAsnArg 1182
Db 3857 GTGTTTCAGATGGCAAGTCAGAGCATGGCACACATCAACAGCTGCTGGCTCAGAAA 3916
QY 1183 AspileTyrPheLysLeuValAsnAlaGlnSer 1193
Db 3917 GGCATCTATTTTCCATGATCAGTGTCCAGGCT 3949
```

Search completed: March 31, 2003, 02:22:27
Job time : 16406.1 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 30, 2003, 01:20:01 ; Search time 1167.91 Seconds
(without alignments)
2304.231 Million cell updates/sec

Title: US-09-873-409-6

Perfect score: 6012

Sequence: 1 MILGTLASLVNGACLPMLP.....QELLNRDIYFKLVNAQSVQ 1195

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -DEV=xlh

-Q/cn2_1/USPTO.spool/US09873409/runat_27032003_115416_19216/app_query.fasta_1.7544
-DB=N_Geneseq_101002 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09873409 @CN1_1_4247 @runat_27032003_115416_19216 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_101002.*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Length	DB	ID	Description
1	5760.5	95.8	3699	24	ABK83223	Human transporter
2	3557	59.2	4279	22	AAD03506	Dog P-glycoprotein
3	3556.5	59.2	4195	22	AAF86128	Cynomolgus monke
4	3554	59.1	3860	21	AAZ49332	Human wild-type mu
5	3554	59.1	3860	24	ABA94365	Human BCRP DNA rel
6	3554	59.1	4279	22	AAD03504	Dog P-glycoprotein
7	3554	59.1	4279	22	AAD03505	Dog P-glycoprotein
8	3549	59.0	4279	22	AAD03488	Dog P-glycoprotein
9	3549	59.0	4317	22	AAD03489	Dog P-glycoprotein
10	3548	59.0	4186	22	AAF86127	Cynomolgus monke
11	3545	59.0	3860	21	AAZ49333	Human G185V mutant
12	3545	59.0	3860	24	ABA94366	Human BCRP DNA rel
13	3545	59.0	4349	22	AAH57442	Human intestine ce
14	3545	59.0	4646	21	AAZ94738	Human ATP binding
15	3545	59.0	4646	24	AAD38994	Human mdrl gene.
16	3545	59.0	4646	24	ABL68592	Kidney cancer rela
17	3545	59.0	4646	24	ABL68880	Kidney cancer rela
18	3545	59.0	5505	17	AAH13394	Hybrid vector pSF-
19	3545	59.0	8630	21	AAZ24041	Retroviral M4 mdr-
20	3545	59.0	8630	21	AAZ24042	Retroviral vector
21	3543.5	58.9	4264	19	AAV66533	Mutated human P-gl
22	3543.5	58.9	4264	19	AAV66534	Mutated human P-gl
23	3543	58.9	4669	8	AAH70752	Sequence of human
24	3540.5	58.9	4425	21	AAZ52048	Rat multidrug resi
25	3539	58.9	4378	11	AAQ04522	Multidrug Resistan
26	3538.5	58.9	4369	21	AAZ52047	Rat multidrug resi
27	3537.5	58.8	3840	24	ABL91687	Human polynucleoti
28	3537.5	58.8	3988	21	AAZ88973	Human MDR-1 DNA.
29	3536	58.8	4669	14	AAQ52726	Sequence of human
30	3534	58.8	4646	15	AAQ72872	Human multidrug re
31	3527	58.7	4669	19	AAV32645	Human P glycoproce
32	3527	58.7	4189	21	AAZ52034	cDNA encoding huma
33	3526.5	58.7	4189	21	AAZ49331	Murine multidrug r
34	3526.5	58.7	4189	24	ABA94367	Mouse BCRP DNA rel
35	3526.5	58.7	4313	14	AAQ38950	Mouse multidrug re
36	3522.5	58.6	4788	21	AAZ49335	Murine multidrug r
37	3522.5	58.6	4788	24	ABA94368	Mouse BCRP DNA rel
38	3517.5	58.5	4233	21	AAZ90198	Rat mdrlb2 (multis
39	3517.5	58.5	4233	22	AAF27498	Rat mdrlb2 multidr
40	3504.5	58.3	3924	21	AAZ94742	Human ATP binding
41	3504.5	58.3	3924	21	AAZ88974	Human MDR-3 DNA
42	3504.5	58.3	3924	24	ABN95801	Gene #2299 used to
43	3476.5	57.8	4254	24	ABK63517	Rat sequence diffe
44	3428.5	57.0	3912	24	ABK63653	Rat sequence diffe
45	2868	47.7	4776	21	AAZ94744	Human ATP binding

ALIGNMENTS

RESULT 1

ABK83223
ID ABK83223 standard; cDNA; 3699 BP.

XX ABK83223;

XX 27-AUG-2002 (first entry)

DE Human transporter and ion channel, TRICH14, Incyte ID 7472030CB1, cDNA.

XX Human; ss; gene; transporter and ion channel; TRICH; transport disorder;
KW neurological disorder; muscle disorder; immunological disorder; cancer;
KW scleroderma; systemic lupus erythematosus; allergy; leukaemia;
KW cell proliferative disorder; cervical cancer; breast cancer;
KW neurodegenerative disorder; Parkinson's disease; Alzheimer's disease;
KW myotonic dystrophy; catatonias; endocrine disorder; diabetes;
KW Grave's disease; gastrointestinal disorder; Crohn's disease;
KW renal disorder; Good pasture's syndrome; viral infection; cirrhosis;

KW bacterial infection; fungal infection; parasitic infection;
 KW protozoal infection; helminthic infection; cardiovascular disorder;
 KW atherosclerosis; hepatic disease.
 XX Homo sapiens.
 XX WO200240541-A2.
 XX 23-MAY-2002.
 XX 25-OCT-2001; 2001WO-US46055.
 XX 27-OCT-2000; 2000US-243989P.
 XX 03-NOV-2000; 2000US-245904P.
 XX 09-NOV-2000; 2000US-247673P.
 XX 17-NOV-2000; 2000US-249661P.
 XX 20-NOV-2000; 2000US-252232P.
 XX 01-DEC-2000; 2000US-250790P.
 XX (INCY-) INCYTE GENOMICS INC.
 XX Tang YT, Yue H, Nguyen DB, Hafalia AJA, Elliott VS, Lu Y;
 PI Wallia NK, Yao MG, Baughn MR, Gandhi AR, Ding L, Sanjanwala M;
 PI Ramkumar J, Arvizu C, Gietzen KJ, Lal PG, Azimzai Y, Khan FA;
 PI Thangavelu K, Thornton M, Lu DAM, Tribouley CW, Warren BA;
 PI Ison CH, Das D, Raumann BE, Policky JL, Kearney L;
 XX WPI; 2002-463570/49.
 DR P-PSDB; ABG61544.
 XX
 XX New transporters and ion channels (TRICH) polypeptides, useful for
 PT diagnosing, preventing, and treating disorders associated with an
 PT abnormal expression or activity of TRICH, e.g. immunological, muscular
 PT or renal disorders
 XX
 XX Claim 5; Page 172; 178pp; English.
 XX
 XX The invention relates to human transporters and ion channels (TRICH)
 CC polypeptides, a naturally occurring amino acid sequence 90 % identical to
 CC TRICH, a biologically active fragment of TRICH or an immunogenic fragment
 CC of TRICH. Also included are an isolated polynucleotide encoding TRICH,
 CC a recombinant polynucleotide comprising a promoter sequence operably
 CC linked to the TRICH polynucleotide, a cell transformed with the
 CC recombinant polynucleotide, a transgenic organism comprising the
 CC recombinant polynucleotide, an isolated antibody that binds specifically
 CC to TRICH, and screening for compounds which bind to TRICH, modulate
 CC TRICH, modulate TRICH expression or are ant/agonists of TRICH.
 CC The polypeptides are useful for diagnosing, treating, and
 CC preventing transport, neurological, muscle, immunological disorders
 CC (e.g. scleroderma, systemic lupus erythematosus, allergies), cell
 CC proliferative disorders such as cancers (e.g. leukemia, cervical or
 CC breast cancers), neurodegenerative disorders (e.g. Parkinson's disease,
 CC Alzheimer's disease), muscular disorders (e.g. myotonic dystrophy,
 CC catatonias), endocrine disorders (e.g. diabetes, Grave's disease),
 CC gastrointestinal disorders (e.g. Crohn's disease), renal disorders
 CC (e.g. Good pasture's syndrome), viral, bacterial, fungal, parasitic,
 CC protozoal and helminthic infections, cardiovascular disorders (e.g.
 CC atherosclerosis), or hepatic diseases (e.g. cirrhosis) and many
 CC other diseases and disorders detailed in the specification. They can also
 CC be used in assessing the effects of exogenous compounds on the
 CC expression of nucleic acid and amino acid sequences of transporters and
 CC ion channels. TRICH or its fragments may also be used in screening for
 CC compounds that specifically bind to and modulate the activity of TRICH.
 CC The polynucleotides can be used to create knock-in humanised animals or
 CC transgenic animals to model human disease. The present sequence
 CC encodes a TRICH protein.
 XX
 XX Sequence 3699 BP; 1116 A; 707 C; 860 G; 1016 T; 0 other;

Alignment Scores:

Pred. No.: 0

Length: 3699

Score: 5760.50

Matches: 1156

Percent Similarity: 95.87%

Conservative: 4

Best Local Similarity: 95.54% Mismatches: 9
 Query Match: 95.82% Indels: 41
 DB: 24 Gaps: 3
 US-09-873-409-6 (1-1195) x ABK83223 (1-3699)
 QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
 DB 145 ATGATCCTGGGTATACCTGACATCAGTGTTCATGGAGCCTGCTTCTTTAATGCCACTG 204
 QY 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
 DB 205 TGTATAGGAGAAATGAGTATACCTTATTAGTGGATGCTAGTCCACACATAACAACA 264
 QY 41 SerPhePheArg-----LeuThr 46
 DB 265 AATTATCAGAACTGTACTCAGTCTCAAGAGAGAGCTGAATGAAGATATGACTCTGTGTGACC 324
 QY 47 LeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeu 66
 DB 325 CTGTATTATGTTGGAATAGGTGTGCTGCTTGTATTTTGGTTACATACAGATTTCCTTG 384
 QY 67 TrpIleIleThrAlaAlaArgGlnThrLysArgIleAArgLysGlnPhePheHisSerVal 86
 DB 385 TGGATTATAACTGCAGCAGCAGACCAAGAGAGATTTCGAAAAACAGTTTTTTCATTTCAGTT 444
 QY 87 LeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGluLeuAsnThrArgMet 106
 DB 445 TTGGCCACAGACATCGCTGGTGTGATAGCTGTGACATCGGTGAACCTTAACACTCGCATG 504
 QY 107 Thr---AspIleAspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGln 125
 DB 505 ACAGATGACATTGACAAATCAGTATGATGATTTGGAGATAAGATTGCTCTGTGTTCATA 564
 QY 126 AsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLysGlyTrpLysLeuThr 145
 DB 565 AACATGCTACTTTTTTCGATTGGCTGGCAGTTGGTTTGGTGAAGGCTGGAAACTCACC 624
 QY 146 LeuValThrIleuSerThrSerProLeuIleMetAlaSerAlaAlaAlaCysSerArgMet 165
 DB 625 CTAGTAGTACTTATCCACAGTCTCTCTTATTAATGGCTTCAGCGCAGCATGTTCTAGATG 684
 QY 166 ValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGlu 185
 DB 685 GTCATCTCTATTGACCACTAAGGAAATTAAGTGCCTATTCCAAAGCTGGGCTGGGAGAA 744
 QY 186 GluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGln 205
 DB 745 GAAGTCTTGTCTCATCCGAAACAGTCATAGCTTTAGGGCCCGAGAGAAAGAACTTCAA 804
 QY 206 ArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSer 225
 DB 805 AGGTATACACAGAAATCTCAAGATGCAAGAGATTTTGGCATAAAAAGGACTATAGCTTCA 864
 QY 226 LysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrp 245
 DB 865 AAGTGTCTCTTGGTCTGTGTACTCTTTATGAATGGAACCTATGACTTGTCTTTTGG 924
 QY 246 TyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAla 265
 DB 925 TATGGAACCTCTTGTATCTTAATGGAGAACCTGGATATACCATCGGGAGCTGTCTTGTCT 984
 QY 266 ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGlu 285
 DB 985 GTTTTCTTTAGTGTAAATCCATAGTAGTTATTCATTTGGAGAGCAGCAGTCCCTCACTTGA 1044
 QY 286 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPro 305
 DB 1045 ACCTTGCAATAGCCCGAGAGAGCTGCTTTCATATTTTCCAGGTATTTATATAGAACC 1104
 QY 306 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 325
 DB 1105 AGTATAGGTAACTTTTCCACAGCTGGATATAAACCTGAATCCATAGAAGAACTGTGGAA 1164

Qy	159	AlaAlaA	CysSerArgMetVal	lleSerLeuThrSer	LySGluLeuSer	AlaTyrSer	178
Db	707	CGCGCCATCTGGC	CAAGATACTATCTT	CATTCTAGTA	AAAGAACTCTTGGC	CTATGCA	766
Qy	179	LysAlaGly	AlaValaGluGlu	ValLeuSerSer	lleArgThrVal	lleAlaPheArg	198
Db	767	AAAGCTGGAGCAGT	AGTCTGAAGAAGTCTTAG	CAGCAATCAGAACTGTGAT	TGCTTTGGATGGC	TTTGGAA	826
Qy	199	AlaGlnGlu	LySGluLeuGln	ArgTyrThrGln	AsnLeuLySGlu	AspAlaLysAsp	218
Db	827	GGACAAAGAA	GAACACTTGA	AGGTACAACAAATTTAGA	GAAGACTAAAGAA	ATTGGG	886
Qy	219	IleLysArg	ThrIleAlaSer	LySValSerLeuGly	AlaValTyrPhe	PheMetAsnGly	238
Db	887	ATAAGAAAGCTAT	CACGGCCCAACATTTCTAT	TGTGTGGCGCTTCTTCT	TATTGATCATCTATGCA	946	
Qy	239	ThrTyrGly	LeuAlaPheTrpTyrGly	ThrSerLeuIleLeu	AsnGlyGluProGlyTyr	258	
Db	947	TCATATGCTCTG	CGCTTCTGGTATGGACCTCCTTGGT	CTCTCCAGTGAA-----TAT	1000		
Qy	259	ThrIleGly	ThrValLeuAlaVal	PhePheSerVal	lleHisSerSerTyrCys	IleGly	278
Db	1001	ACTATTGGACA	AGTACTCACTGCTCTTCTTTCTG	TATTAAATTGGCGCTTTTAGT	ATTGGA	1060	
Qy	279	AlaAlaVal	ProHisPheGluThrPhe	AlaIleAlaArgGly	AlaAlaPheHisIle	Phe	298
Db	1061	CAGCATCTCCCA	AGCATTTGAACATTTGCAACGCAAG	GAGCAGCTTATGAAATCTTC	1120		
Qy	299	GlnValIle	AspLySlySProSerIle	AspAsnPheSerThr	AlaGlyTyrLysProGlu	318	
Db	1121	AGNATATTGCA	ATTAACACAGCATTTGACACTATTG	CAAGATGGACATTAACAGAT	1180		
Qy	319	SerIleGluGly	ThrValGluPhe	LySAsnValSerPhe	AsnTyrProSerArgProSer	338	
Db	1181	AAATATTAAAGG	CAAAATTTGGAAATTC	AAAAATGTTCACTTCACTT	TGACCTTCTCGAAAAAGAA	1240	
Qy	339	IleIysIle	LeuLySGlyLeu	AsnLeuArgIleLys	SerGlyGluThrValAlaLeuVal	358	
Db	1241	GTTAAGATCTTAA	AGGGTCTCAACCTGAAGGTTCA	GATGGGACAGATGGCGCTGGT	1300		
Qy	359	GlyLeu	AsnGlySerGlyLys	SerThrValValGlnLeu	LeuGlnArgLeuTyrAspPro	378	
Db	1301	GGGACAGTGGCT	CGGGAAGACGACCGCTGCAGCTGTATG	CAGAGGCCTATGACCCC	1360		
Qy	379	AspAspGly	PheIleMetVal	AspGluAsnAspIle	ArgAlaLeuAsnValArgHisTyr	398	
Db	1361	ACAGATGGCAT	TGCTGTATTGATGCACAGGACATTAG	GCACCAATAATGTAAGGCATCTT	1420		
Qy	399	ArgAspHis	IleGlyValValSerGln	GluProValLeuPheGlyThr	ThrIleSerAsn	418	
Db	1421	CGGGAAATTA	CTCGTGTGTGAGTCAGGAGCTGTGTGT	TTGCCACACAGATGACTGAA	1480		
Qy	419	AsnIleLys	TyrGlyArgAsp	AspValThrAspGluGluMetGlu	ArgAlaAlaArgGlu	438	
Db	1481	AACATTGCCTAT	GGCCGCAAAATGTCACCATGGATGAGAT	TGAGAAAGCTGTTAAGAA	1540		
Qy	439	AlaAsnAla	TyrAspPheIleMetGlu	PheProAsnLySGlu	AsnThrLeuValGlyGlu	458	
Db	1541	GCCAAATGCC	ATTATGATTTATCATGAAATACCT	TAATAAATTTGACACTCTG	GGTGGAGAG	1600	
Qy	459	LysGlyAla	GlnMetSerGlyGlyGln	LysGlnArgIleAlaIleAlaArgAlaLeuVal	478		
Db	1601	AGAGGGCC	CAGCTGAGTGTGTGACAGAAACAGAAAT	CGCCATTGCTCGGGCCCTG	1660		
Qy	479	ArgAsnPro	LysIleLeuIleLeu	AspGluAlaThrSerAlaLeu	AspSerGluSerLys	498	
Db	1661	CGCAACCCCA	AGATTTCTTCTGCTGGATGAGGCAACG	TCAGCTCTGGACACTGAAAGTAA	1720		
Qy	499	SerAlaVal	GlnAlaLeuGluLysAla	SerLySGlyArgThrThrIleValAla	518		
Db	1721	GCAGTGTTCAG	TGTGGCTTGGATAGGCCACAGAAAGC	CGGACTTACCATGTGATAGCT	1780		

Qy	519	HisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeu	538
Db	1781	CATCGTTGTGTACAGTTCGTAAATGCCGATGTCATTCGTGTTGTGATGATGGAGTCATT	1840
Qy	539	AlaGluLysGlyAlaHisAlaGluLeuMetAlaLysAcGlyLeuTyrTyrSerLeuVal	558
Db	1841	GTGGAGAAAGGAATCATGATGACATCATGAAGACAGAGGCACTTTACTTCCAAACTTGTCT	1900
Qy	559	MetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu	578
Db	1901	ACAATGCAG--ACAAGAGAAATGAATGAGTTAGTAAATGCCACTGGTGAATCCAAAC	1957
Qy	579	ArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle---	597
Db	1958	AGTGAAGAGTCATCCCTGGAAATG--TCTCCAAAAGAGTTCCAGGTCACAGTTTAAATAAA	2014
Qy	598	-----AspLysAlaGluGluSer	603
Db	2015	AGAAGATCAACTCGCAGGAGTATACATGCACACCAAGGCCAAGACAGAACCTTGGTACA	2074
Qy	604	ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn	623
Db	2075	AAAGAGGACTTGAATCAGAAATGACCTCCAGTTTCTCTTCGGAGGATTTCTGAAGCTGAAC	2134
Qy	624	LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal	643
Db	2135	TCAACTGAATGGCGCTTATTTGTGTGGTATATTTGTGCTATTATATAACCGGAGGCGTC	2194
Qy	644	HisProValPheSerIleIlePheAlaLysIleIleThrMetPhe--GlyAsnAsnAsp	662
Db	2195	CAACGACGATTTTCAATAATATTTTCAAGGATATATAGGATCTTTTACCAGAGATGAGGAT	2254
Qy	663	LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal	682
Db	2255	CCTGAACAAACACGACAGAATAGTAACATGTTTCTGTATTGTTCTTAGCTCTTGGAAAT	2314
Qy	683	IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu	702
Db	2315	ATTTCTTTTATTAACATTTTCTCCAGGCTTTCACATTTGGCAAGAGCTGGGGAGATCCTC	2374
Qy	703	ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe	722
Db	2375	ACTAAGCGGCTTCGATACATGTTTTCAGATCCATGCTGACAGAGATGTCAAGCTGGTTT	2434
Qy	723	AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln	742
Db	2435	GATGACCTTAAACACACCCTGGAGCATTTGACACACGAGCTTGCAATGATCGCGCTCA	2494
Qy	743	IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly	762
Db	2495	GTTAAGGGGCTATAGTTCAGGCTTGCTGTCTCATTCACAGATATAGCAAACTTTGGG	2554
Qy	763	LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle	782
Db	2555	ACAGGCATTAATATATCCCTTAATCTATGTTGGCAATTAACACTTTTACTCTTAGCAATT	2614
Qy	783	AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn	802
Db	2615	GTACCCATCATTCGATAGCAGAGTGTGTAATGAAAAATGTTGCTGGACAGCACTG	2674
Qy	803	LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle	822
Db	2675	AAAGATTAAGAAAGAGCTAGAAGAGAGCTGGGAAGATTGCTACAGAGCCATCGAAACTTC	2734
Qy	823	ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu	842
Db	2735	CGAAGCTGTGTTCTTTTGACTCGGGAGCAGAGTTTGAATATCATGTATGTCACAGAGTTTG	2794
Qy	843	GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe	862
Db	2795	CAAGTACCATCAGAAACTCTTTGAGGAAGACACACATCTCCGGGCTCTCATTTCTTATC	2854
Qy	863	SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle	882

[illegible]

1384 GTTGGAAACACGGCGTGTGGGAAGAGCACACCGTTCAGCTGTATGCAGAGCGCTTTATGAC 1444
 QY ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis 397
 1444 CCCACAGAGGCGTGTCACTGTTGATGGACAGGATATTAGGACCATAAACGTAAGGTTT 1503
 QY TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSer 417
 1504 CTACGGGAAATCATCGGTGTGGTGATCATGGAACCTGATTGTTTGGCCACACGATAGCT 1563
 QY AsnAsnIleLysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArg 437
 1564 GAAACAATTCGCTATGGTCTGAGATGTCACCATGGATGGATGGATGGAGAAGCTGTCAAG 1623
 QY GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly 457
 1624 GAAGCCAAATGCCTATGACTTTATCATGAAACTGCCTCAGAAATTTGACACCCCTGGTTGA 1683
 QY GluLysGlyAlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu 477
 1684 GAGAGAGGGGCCACGCTGAGTGTGGCGAAGACAGAGAGGATCGCCATTGTCAGCTGCCCTG 1743
 QY ValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 497
 1744 GTTCGCAACCCCAAGATCTCTCTGTCGACGAGGCCAGCTCAGCCCTGGACAGAAAGT 1803
 QY LysSerAlaValGlnAlaLeuGluLysAlaSerLysGlyArgThrThrIleValVal 517
 1804 GAACAGCTGGTTCAAGTGGCTCTGGNATAGCGCAGAAAGGTTCGACACCATTTGTCATA 1863
 QY AlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMet 537
 1864 GCTCATCGTTGTCTACGGTTGCTAATCGCAGCTCATCGCTGGTTTCGATGATGGAGTC 1923
 QY LeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeu 557
 1924 ATTGTGAGAAAGAAATCATGATGAGCTCATGAAGAAGAGGCAATTTACTTCAAACTT 1983
 QY ValMetSerGlnAspIle-----LysLysAlaAspGlu----- 568
 1984 GTCACAATGCAGACAGCAGGAAATGAAATTGAATAGAAATGCAGCTGATGAATCCAA 2043
 QY ---GlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeu----- 584
 2044 AGTGAAATTGATACCTTGGAAATGCTCTTCACATGATTTCAGGATCCAGCTTAATAAGAAAA 2103
 QY ---ProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGlu 602
 2104 AGATCCACTCGTAGGAGTGTCCGTGGATCACAAGGCCAA-----GACAAAGACTTAGT 2157
 QY SerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeu 622
 2158 ACCAAAGAGCTCTGGATGAAAGTATACCTCCAGTTTCTTTTGGAGGATTATGAGCTA 2217
 QY AsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThr 642
 2218 AATTAACTCAGTGGCCTATTGTTGTGTGTATTGTCGTCATTATATAAATGGAGGT 2277
 QY ValHisProValPheSerIleIlePheAlaLysIleIleThrMetPhe---GlyAsnAsn 661
 2278 CTGCAACACGACATTCGCGTAATATTTTCAAAGATATTATAGGGATTTTTTACAAGAAATGAT 2337
 QY AspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGly 681
 2338 GATCCGAAACAAACACGACGAATAGTAATGTTTTCACATTATGTTTCTAGTCTTGGGA 2397
 QY ValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyLeuIle 701
 2398 ATTGTTCTTTTATACATTTTTTCTTCAGGCGCTTCACATTTGGCAAAGCTGGAGATC 2457
 QY LeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrp 721
 2458 CTCACCAAGGGGTCCGATACATGTTTCCGATGTTTCCGATCCATGCTCAGACAGATGTGAGCTG 2517

QY 722 PheAspGluLysGluAenSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAla 741
 DB 2518 TTTGATGACCCCTAAACACACCATCGGAGCATTCAGTACAGGCTCGCCATGATGCTGCT 2577
 QY 742 GlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAenAlaThrAenMet 761
 DB 2578 CAAGTTAAAGGGCTATAGGTTCCAGGCTTGCTATTAATACCCAGAAATATAGCAAAATCTT 2637
 QY 762 GlyLeuSerValIleIleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSer 781
 DB 2638 GGGACAGGAATAATATATCTTAATCTATGTTGGCACTGACATCTGTACTCTTATGCA 2697
 QY 782 IleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAla 801
 DB 2698 ATTGTACCCATCATTCGAATAGCAGGAGTGTGTAATGAATGAAATGTTGCTGGCAAGCA 2757
 QY 802 AsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAen 821
 DB 2758 CTGAAAGATAAGAAAGAACTAGAGGTCCTGGGAAGTCGCTACTGAAAGCAATAGAAAAC 2817
 QY 822 IleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMet 841
 DB 2818 TTCGGAACCTGTTCTTTGACTCAGGAGCAGAAAGTTTGAACATATGTATGATCAGAGT 2877
 QY 842 LeuGlnThrGlnHisArgAenThrSerLysLysAlaGlnIleIleGlySerCysTyrAla 861
 DB 2878 TTCAGGTACCATACAGAACTCTTTGAGGAAGACACACATCTTTGGAATCAGTTTTC 2937
 QY 862 PheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeu 881
 DB 2938 TTCACGAGGCAATGATGATATTTTCTATGCTGGATGTTTCGGTTTGAGGCTACTTG 2997
 QY 882 IleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyr 901
 DB 2998 GTGGCACATAGTCTCATGACCTTTGAGGATGTTCTGTGTATATTTTCAGCTGTGCTTT 3057
 QY 902 GlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSer 921
 DB 3058 GTGTGCCATGCGCGGGGCAAGTCAGTTCTATGCTCTGCTGACTATGCCAAGCAAGTA 3117
 QY 922 GlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysPheAsnIleAspSerArgSerGln 941
 DB 3118 TCAGCAGCCCATCATCATCATGATCAATGAAAGAAACCCCTTGTGATGACACTACAGACA 3177
 QY 942 GluGlyLysLysProAspThrCysGluLysAsnLeuGluPheArgGluValSerPhePhe 961
 DB 3178 GAAGGCGCTAAAGCGCAACATTTGGAAGGAAATGTCTACATTTAATGAAGTTGTATTCAAC 3237
 QY 962 TyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGly 981
 DB 3238 TATCCCAACCCGACTGGACATCCCGAGTCTTCAGGGGCTGAGCCTGGAGTGAAGAGGGC 3297
 QY 982 LysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeu 1001
 DB 3298 CAGAGCTGCGCCCTGCTGGGCGAGCAGTGGCTGTGGGAAGACACGCGTGTCCAGCTCCTG 3357
 QY 1002 GlnArgLeuTyrAspProValGlnGlyValLeuPheAspGlyValAspAlaLysGlu 1021
 DB 3358 GAGCGGTCTATGACCCCTTGGCGGGGAAAGTGTCTTGACGGCAAGAAATAAAGCAA 3417
 QY 1022 LeuAenValGlnTyrLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPhe 1041
 DB 3418 CTGAATGTTTCAGTGGCTCCGAGCACACTGGGCATCTGTGCCAGGAGCCATCTCTGTTT 3477
 QY 1042 AsnCysSerIleAlaGluAenIleAlaTyrClyAspAsnSerArgValValProLeuAsp 1061
 DB 3478 GACTGCGAGCATTAGTGAACAATTTGCTATGGAGACAACAGCCGGGTGTGTCCACAGGAA 3537
 QY 1062 GluIleLysGluAlaAlaAenAlaAlaAsnIleHisSerPheIleGluGlyLeuProGlu 1081
 DB 3538 GAGATCGTGGGCGAGCCAAAGGAGGCCAATATACACGCTTTCATCGAGTCACTGCCTAAT 3597

QY 1082 LysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArg 1101
 DB 3598 AAATATAGCACAGAGTAGGACAAAGAACTCAGCTCTCTGGTGGCCAGAAACAACGC 3657
 QY 1102 LeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThr 1121
 DB 3658 ATTGCCATAGCTCGTGCCCTTGTAGACAGCCTCATATTTCTTTGGATGAGCCACA 3717
 QY 1122 SerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThr 1141
 DB 3718 TCAGCTCTGGATACAGAAAGTGAAGGTTGTCGAAGAGCCCTGGACAAAGCCAGAA 3777
 QY 1142 GlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIle 1161
 DB 3778 GGCGGTACCTGCTGATTTGCTCATCCGCTGTCCACCATCAGAATGCAGACTTAATA 3837
 QY 1162 ValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsn 1181
 DB 3838 GTGGTGTTCAGAATGGCAGAGTCAAGGAGCAGCGCACATCAGCAGCTGCTGGCACAG 3897
 QY 1182 ArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1193
 DB 3898 AAAGCATCTATTTTCAATGCTGCTCAGTGTCCAGGCT 3933
 RESULT 4
 AA249332
 ID AA249332 standard; cDNA; 3860 BP.
 XX AA249332;
 AC AA249332;
 XX 14-MAR-2000 (first entry)
 DT
 DE Human wild-type multidrug resistance-1 (MDR-1) cDNA.
 KW Multidrug resistance; MDR-1; P-glycoprotein;
 KW transmembrane efflux pump; haematopoietic stem cell; transduction;
 KW bone marrow transplantation; chemotherapy; radiation therapy; cancer;
 KW gene therapy; gene replacement; genetic defect; thalassaemia;
 KW Gaucher's disease; sickle cell anaemia; leukaemia; ex vivo expansion;
 KW cytokine; wild-type; ds.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 1..3843
 FT /tag= a
 FT /product= "Human wild-type MDR-1 protein"
 FT mutation replace (553..555, GTT)
 FT /tag= b
 FT /note= "cDNA sequence of G185V human mutant MDR-1 given
 in AA249333"
 PN WO9961589-A2.
 XX
 XX 02-DEC-1999.
 PD
 XX
 PF 27-MAY-1999; 99WO-US11825.
 XX
 PR 28-MAY-1998; 98US-0086988.
 XX
 PA (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
 XX
 XX Sorrentino B, Bunting K;
 PI
 XX WPI; 2000-072615/06.
 DR P-PSDB; AAY58186.
 XX
 PT Ex vivo expansion of hematopoietic stem cells transduced with a
 PT sequence encoding human multidrug resistance-1, used for bone marrow
 PT transplantation -
 XX
 PS Claim 10; Page 68-70; 113pp; English.

CC This sequence represents cDNA encoding human wild-type
 CC multidrug resistance protein MDR-1. MDR-1 is a transmembrane
 CC efflux pump, responsible for the export of drugs from cells,
 CC particularly cancer cells. Wild-type MDR-1 shows increased
 CC resistance to etoposide and decreased resistance to vinca
 CC alkaloids compared with a mutant form (AY58187) where the Gly at
 CC position 185 is replaced by Val. The invention relates to transducing
 CC haematopoietic stem cells with nucleic acid encoding an MDR protein
 CC and culturing the modified cells. The modified haematopoietic stem
 CC cells are useful in bone marrow transplantation (to reconstitute
 CC haematopoietic systems in patients who have undergone chemotherapy or
 CC radiation therapy) and in ex vivo gene therapy of genetic defects in
 CC cells derived from haematopoietic stem cells, e.g., thalassemia,
 CC Gaucher's disease, sickle cell anaemia or leukaemia. The modified
 CC cells can also be used to identify factors involved in regulating
 CC proliferation and differentiation in haematopoietic stem cells.
 CC Haematopoietic stem cells that express MDR-1 will be protected against
 CC chemotherapeutic agents, so can be engrafted while the patient is
 CC undergoing chemotherapy. Expansion of (rare) haematopoietic stem cells
 CC provides sufficient cells to permit standard biochemical analysis.
 CC Overexpression of MDR-1 allows cytokine-driven expansion of
 CC haematopoietic stem cells by at least 10-fold compared with a maximum
 CC of 4-fold in known procedures.

XX
 SQ Sequence 3860 BP; 1135 A; 746 C; 957 G; 1022 T; 0 other;

Alignment Scores:

Pred. No.: 0 Length: 3860
 Score: 3554.00 Matches: 682
 Percent Similarity: 75.63% Conservative: 246
 Best Local Similarity: 55.58% Mismatches: 263
 Query Match: 59.12% Indels: 36
 DB: 21 Gaps: 9

US-09-873-409-6 (1-1195) x AA249332 (1-3860)

Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
 Db 151 ATGGTGGTGGAACTTTGGTGCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 210
 Qy 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
 Db 211 GTGTTGGAGAAATGACAGATATCTTTGCAAAATGCAGGAAATTTAGAAAGATCTGATGTC 270
 Qy 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
 Db 271 AACATCACTAATAGAGATGATATCATATGATACAGGGTTCTTCATGATCTGGAGGAAGAC 330
 Qy 44 -----ArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
 Db 331 ATGACCAGATATCCCTATTATTACAGTGAATTTGGTGTGGGTGCTGCTTAC 390
 Qy 62 IleGlnIleSerLeuTrpIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
 Db 391 ATTCAAGTTTCATTTTGGTGGCTGGCAGCTGGAAGACAAATACACAAATTTAGAAACAG 450
 Qy 82 PhePheHieSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101
 Db 451 TTTTTCATGCTATATACGACAGAGATAGGCTGGTTGATGTCACGATGTTGGGAG 510
 Qy 102 LeuAsnThrArgMetThr-----AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
 Db 511 CTTAACACCCGCTACAGATGATGCTCTTAAGATTAATGAAGGTATTGGTGACAAAT 570
 Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
 Db 571 GGAATGTCTTTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGGATTTTACACGT 630
 Qy 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
 Db 631 GGTGGAAGCTAACCTTTGATTTTGGCCATCATGCTCTTCTTGGACTGTGCTGCTGCT 690
 Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180

Db 691 GTCTGGGCAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATCAAAAGCT 750
 Qy 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
 Db 751 GGAGCAGTAGCTGAAGAGGTCTGGCAGCAATTAGAAGCTGTGATGTCATTTGGAGACAA 810
 Qy 201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
 Db 811 AAGAAGAACTTTGAAAGGTACAACAAAATTTAGAAGAGCTAAAGAAATTTGGCATAAAG 870
 Qy 221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
 Db 871 AAAGCTATTACAGCCCAATATTCTATAGGTGCTGCTTCTCTGCTGATCATCTTAT 930
 Qy 241 GlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
 Db 931 GCTCTGGCCCTCTGGTATGGGACCACTTGGTCTCTCAGGGGAA-----TATCTATT 984
 Qy 261 GlyThrValLeuAlaValPhePheSerValIleHieSerSerTyrCysIleGlyAlaAla 280
 Db 985 CGACAAGTACTCACTGATTTCTTTCTGTATTAAATGGGGCTTTTAGTGTGACACAGGCA 1044
 Qy 281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHieIlePheGlnVal 300
 Db 1045 TCTCCAAGCATTTGAAGCATTTGCCAAATGCAGAGGAGCAGCTTATGAAATCTTCAAGATA 1104
 Qy 301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
 Db 1105 ATGTATATTAAGCAAGTATTGACAGCTATTGAGAGTGGGCACAAACAGATATATT 1164
 Qy 321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
 Db 1165 AAGGGAAATTTGGAATTCAGAAATGTTTCACTTCACTTCACTTCACTTCACTTCACTTCA 1224
 Qy 341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
 Db 1225 ATCTTGAAGGGCTGAACCTGAAGTGCAGAGTGGGCAGACGGTGGCCCTGGTGGTGAAC 1284
 Qy 361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
 Db 1285 AGTGGCTGTGGGAAGAGCACACAGTCCAGCTGTATGACAGGCTCTATGACCCACAGAG 1344
 Qy 381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
 Db 1345 GGGATGTCAGTGTGATGACAGAGATATTAGGACCATTAATGTAAGTTTCTTACGGGAA 1404
 Qy 401 HisIleGlyValValSerGlnProValLeuPheGlyThrThrIleSerAsnAsnIle 420
 Db 1405 ATCATTTGGTGTGTGAGTCAAGAACCTGTATTGTTTCCACCACCATGCTGTAACCAAT 1464
 Qy 421 LysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
 Db 1465 CGCTATGGCCGTAATAATGTCCATGGATGAGATTTGAGAAAGCTGTCAAGGAAGCAAT 1524
 Qy 441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyLysGly 460
 Db 1525 GCCTATGACTTTTATCATGAACCTGCTCATATAATTTGACACCCCTGTTGGAGAGAGGG 1584
 Qy 461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
 Db 1585 GCCCAGTTGAGTGTGGGAGAGCAGAGGATCGCCATTTGACACCCCTGCTGCTGCTGCTGCT 1644
 Qy 481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
 Db 1645 CCCAAGTCTCTGCTGGATGAGCCAGCTGACCTTGGACACAGAAACGAGCAGGTG 1704
 Qy 501 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 520
 Db 1705 GTTCAGTGGCTCTGATAGGCCAGAAAGGTGGACCAACCATTTGATGATGCTCATCGT 1764
 Qy 521 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 540

Db 1765 TTGCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTCGATGATGAGTCATGTTGGAG 1824
Qy 541 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrSerLeuValMetSer 560
Db 1825 AAGGAAATCATGATGACTCATGAAGAAAGGACATTTACTTCAAACTGTGCACAATG 1884
Qy 561 Gln-----AspIleLysLysAlaAspGlnMetGlySerMetThr 574
Db 1885 CAGACAGCAGGAAATGAAGTTGAATTAGAAATGACGCTGATGAATCCAAAGAGTGAAT 1944
Qy 575 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal----- 589
Db 1945 GATGCGCTTGGAATGCTTCAAAATGATTCAGATCCAGTCTAATAAGAAAGATCAACT 2004
Qy 590 ---LysSerIleLys---SerAspPheIleAspLysAlaGluSerThrGlnSerLys 607
Db 2005 CGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACCAAGAGGCTCTG 2064
Qy 608 GluIleSerLeuProGluValSerLeuLysIleLeuLysLeuAsnLysProGluTyr 627
Db 2065 GATGAAAGTATACCTCCAGTTCTCTTTGGAGGATTATGAAGCTAAATTTAACTGAATG 2124
Qy 628 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 647
Db 2125 CCTTATTTTGTGGTGTATTTTGTGCCATTTATTAATGGAGCGCTGCACAACGACATTT 2184
Qy 648 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 666
Db 2185 GCAATATATTTTCAAGATTATAGGGTTTTTACAGAAATGATGATCCTGAAACAAAA 2244
Qy 667 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 686
Db 2245 CGACAGAAATAGTAATGTTTTCACATATGTTCTAGCCCTTGGAAATATTTCTTTTATT 2304
Qy 687 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 706
Db 2305 ACATTTTTCCTTCAAGGTTTTACATTTTGGCAAAAGCTGGAGAGATCCTCAACAAAGCGGCTC 2364
Qy 707 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaIlePheAspGluLysGlu 726
Db 2365 CGATACATGTTTCCGATCCATGCTCAGACAGATGTGAGTTGGTTTGATGACCCTAAA 2424
Qy 727 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 746
Db 2425 AACACCACTGGACATTTGACTACCGCTCGCCAATGATGCTCAAGTTAAAGGGCT 2484
Qy 747 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 766
Db 2485 ATAGGTTCCAGGCTTGCTGTAATATACCAGATATATAGCAAAATCTTGGGACAGGAAT 2544
Qy 767 IleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeu 786
Db 2545 ATATCTTTCATCTATGTTGGCAACTAACACTGTACTCTTAGCAATTTGACCCATCAT 2604
Qy 787 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 806
Db 2605 GCAATAGCAGAGTGTGTAATGAAATCTTGTCTGCACAGCACTGAAAGATAAGAA 2664
Qy 807 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 826
Db 2665 GAACTAGAAGGTGCTGGGAAGATCGCTACTTGAAGCAATAGAAAACCTTCCGAACCGTTGTT 2724
Qy 827 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 846
Db 2725 TCCTTGACTCAGGACGCAAGATTGTAACATATATGATGCTCGAGATTTCAGAGTACCATAC 2784
Qy 847 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 866
Db 2785 AGAACTCTTTGAGAAAGCACACATCTTTTGAATATACATTTTCTTCCACCCAGGCAATG 2844
Qy 867 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrIleuIleGlnAlaGlyArg 886
Db 2845 ATGTATTTTTCCTATGCTGATGTTTCCGGTTTGGAGCTTACTTGTGGGCACATAAATC 2904

Qy 887 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 906
Db 2905 ATGAGCTTTGAGAGTCTCTGTTAGTATTTTTCAGCTGTTGTTCTTGGTCCATGGCCGTG 2964
Qy 907 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaHisLeu 926
Db 2965 GGCAAGTCAAGTTCATTTGCTCTGCTATGCAAGCCAAATATATCAGCAGCCACATC 3024
Qy 927 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluLysLysPro 946
Db 3025 ATCATGATCATTTGAAAAAACCCCTTTGATTGACGCTACAGCACGGAAGCCCTAATGCCG 3084
Qy 947 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 966
Db 3085 AACACATTCGGAAGAAATGTCACATTTGTTGTAAGTTGTTTCACTATCCACCCGACCG 3144
Qy 967 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 986
Db 3145 GACATCCCACTGCTTCAGGCACTGAGCCTGGAGGTGAAGAGGCCAGACGCTGGCTCTG 3204
Qy 987 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 1006
Db 3205 GTGGGACAGTGGCTGTGGAGAGACACATGTTCCAGCTCTGGAGCGGTTCTTACGAC 3264
Qy 1007 ProValGlnGlyGlnValLeuPheAspGlyValAlaLysAlaLysGluLeuValGlnTyr 1026
Db 3265 CCCTTGGCAGGAAAGTGTCTGTGTCGCAAGAAATAAAGCAGCTGAATGTTCACTGG 3324
Qy 1027 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 1046
Db 3325 CTCGAGACACACTGGGCATCGTCCAGAGAGCCCATCTCTGTTGACTGCACATTTGCT 3384
Qy 1047 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluLeuLysGluAla 1066
Db 3385 GAGAACATTCCTATGGAGACCAACAGCGGGTGTCTCAGGAAGAGATCGTGGAGGCA 3444
Qy 1067 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 1086
Db 3445 GCAAAGGAGGCCCAACATACATCGCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAA 3504
Qy 1087 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 1106
Db 3505 GTAGAGACAAAGGAACCTCAGCTCTCTGTGGCCAGAAACAACGATTCCTAGCTCGT 3564
Qy 1107 AlaLeuLeuGlnLysProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspAsn 1126
Db 3565 GCCTTGTGTAGACAGCTCATATTTGCTTTGGATGAAGCCACGTCAGCTCTGGATACA 3624
Qy 1127 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1146
Db 3625 GAAAGTGAAGAGTTGTCCAAGAGCCCTGGACAAAGCCAGAGAGCCGCGCCTGCATT 3684
Qy 1147 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1166
Db 3685 GTGATTGCTCACCGCTGTCCACCATCCAGATTCAGACTTAATAGTGTGTTCAGAAT 3744
Qy 1167 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1186
Db 3745 GGCAGAGTCAAGGAGCATGGCAGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTT 3804
Qy 1187 LysLeuValAsnAlaGlnSer 1193
Db 3805 TCAATGCTCAGTCTCCAGCT 3825
RESULT 5
ABA94365
ID ABA94365 standard; DNA; 3860 BP.
XX
AC ABA94365;
XX 26-MAR-2002 (first entry)
DT
XX

DE Human BCRP DNA related seq Id No. 1.

XX Stem cell; ATP transport protein; ATP-binding cassette; antiparkinsonian;
 KW hepatotropic; neurodegenerative; cytostatic; antianemic; muscular; BCRP;
 KW cardiant; gene therapy; ds.
 XX

OS Homo sapiens.

XX Key Location/Qualifiers
 FT CDS 1..3843
 FT /*tag= a

PN WO200192877-A2.

XX PD 06-DEC-2001.

XX PF 30-MAY-2001; 2001WO-US17459.

XX PR 31-MAY-2000; 2000US-0584586.

XX PR 29-MAY-2001; 2001US-0866866.

XX PA (SUJUD-) ST JUDE CHILDREN'S RES HOSPITAL.

XX PI Sorrentino B, Schuetz J;

XX DR WPI; 2002-114368/15.

XX DR P-PSDB; ABB07266.

XX PT Identifying a stem cell, for treating e.g., muscular dystrophy,
 PT myocardial infarction, Parkinson's disease, or neurodegenerative
 PT disorders, comprises detecting the expression of an ATP transport
 PT protein (BCRP) by a cell -

XX PS Disclosure; Page 53-55; 87pp; English.

XX CC The invention provides a method of identifying and/or isolating a stem
 CC cell that involves detecting the expression of an ATP transport protein
 CC containing a conserved ATP-binding cassette (BCRP) by a cell in a sample
 CC comprising stem cells. The isolated stem cells may be used in the
 CC treatment of diseases such as muscular dystrophy, degenerative liver
 CC disorders, myocardial infarction, Parkinson's disease, degenerative
 CC disorders of the brain, and for tissue regeneration or replacement.
 CC Haematopoietic cells can be used in bone marrow transplants (e.g., for
 CC treatment of leukemia) and for ex vivo gene therapy for treating blood
 CC diseases such as sickle cell anemia and thalassemia. The stem cells can
 CC also be used as cell targets in gene therapy protocols. The present
 CC sequence represents a sequence related to the BCRP for which no relevant
 CC information has been provided in the specification.

XX SQ Sequence 3860 BP; 1135 A; 746 C; 957 G; 1022 T; 0 other;

Alignment Scores:

Pred. No.:	0	Length:	3860
Score:	3554.00	Matches:	682
Percent Similarity:	75.63%	Conservative:	246
Best Local Similarity:	55.58%	Mismatches:	263
Query Match:	59.12%	Indels:	36
DB:	24	Gaps:	9

US-09-873-409-6 (1-1195) x ABA94365 (1-3860)

Qy 1 MetIleLeuGlyLeuAlaSerLeuValAsnGlyValaCysLeuProLeuMetProLeu 20

Db 151 ATGGTGGTGGGAACTTGGCTGCATCATCCATGGGGCTGGAGCTTCTCTCATGATGCTG 210

Qy 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31

Db 211 GTGTTTGGGAAATGACAGATATCTTTGCAATGACAGGAAATTTAGACATCTGATGTCA 270

Qy 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43

Db 271 AACATCACTAATAGAGATGATATCAATGATACAGGGGTCTTCTATGAATCTGGAGGAAGAC 330

Qy 44 -----ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaIlePheGlyTyr 61
 Db 331 ATGACCATATATCCCTATTTATACAGTGAATTTGGTGGGGTCTGCTGCTCTTAC 390
 Qy 62 IleGlnIleSerLeuTriPheIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
 Db 391 ATTGAGTTTCATTTTGGTGGCTGGCAGCTGGAAGACAAATACACAAATTTAGAAACAG 450
 Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101
 Db 451 TTTTTCATGCTATAATCGCAGCAGGATAGGCTGTTGATGTGCACGATGTGGGAG 510
 Qy 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
 Db 511 CTTAACACCCGACTTACAGATGATGCTCTAAGATTAATGAAGGTATTGGTGACAAATTT 570
 Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
 Db 571 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGGATTACAGT 630
 Qy 141 GlyTriPlysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
 Db 631 GGTTCGAAGCTAACCTTGTGATTTTGGCCATCAGTCTCTGTTCTTGGACTGTGCTGCT 690
 Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
 Db 691 GTCTGGCAAGATATCTATCTTCTGATAAAGAACTCTTAGCGTATGCAAAAGCT 750
 Qy 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
 Db 751 GGAGCAGTAGCTGAAGAGCTCTTGGCAGCAATTAGAATGATGATTCATTTGGAGACAA 810
 Qy 201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
 Db 811 AAGAAGAACTTGAAGGTACAAACAAATTTAGAAGAGCTAAGAAGATTGGATAAG 870
 Qy 221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
 Db 871 AAGCTATTATACAGCAATATTCTATAGTGTGCTGCTCTGCTGCTATCATCATCTTAT 930
 Qy 241 GlyLeuAlaPheTriPtyGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
 Db 931 GCTTGGGCTTCTGTAAGGACCACTTGGTCTCTCAGGGAA-----TATTCTATT 984
 Qy 261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
 Db 985 GGACAGTACTCAGTATTCTTCTGTATTAAATGGGGCTTTAGTCTGGACAGGCA 1044
 Qy 281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
 Db 1045 TCTCCAAGCATTGAAGCATTTGCAAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1104
 Qy 301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
 Db 1105 ATTGATAATAAGCCAAAGTATTGACAGCTATTGCAAGAGTGGGCAACAAACAGATAATT 1164
 Qy 321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
 Db 1165 AAGGGAATTTGGAATTCAGAAATGTTTCACTTCAGTTACCCATCTCGAAAGAGTTAG 1224
 Qy 341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
 Db 1225 ATCTTGAAGGCTGAACCTGAAGGTGCAGAGTGGGACAGCGTGGCCCTGTTGGAAAC 1284
 Qy 361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
 Db 1285 AGTGCGTGTGGGAAGAGCACACAGTCAGCTGATGCGAGGCTCTATGACCCCAAGAG 1344
 Qy 381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
 Db 1345 GGGATGGTGCAGTGTGATGGAGCAGGATATTAGACCATTAATGTAAAGTTTCTACGGGAA 1404
 Qy 401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnIle 420

QY	259	ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly	278
DB	1001	TCTATTGGACAAGTACTCACTGCTCTCTTTCTGTATTAAATTTAGGTGGCTTTTAGTATTGGA	1060
QY	279	AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe	298
DB	1061	CAGGCATCCCCAAGCATTGAAGCATTTGCAACGCAAGAGAGAGAGTATTAGAAATCTTC	1120
QY	299	GlnValIleAspLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu	318
DB	1121	AGATTAATTGACAAATAAACCAAGCATTTGACAGCTATTCCGAAGGTGGACATAAACAGAT	1180
QY	319	SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer	338
DB	1181	AATATTAAAGGAAATTTGGAATTCAAAATGTTCATCTTCACTTACCCTTCTCGAAAGAA	1240
QY	339	IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal	358
DB	1241	GTTAAGATCTTAAAGGCTCTCAACCTGAAGGTTCAAGTGGCCACAGACAGTGGCGCTGGT	1300
QY	359	GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro	378
DB	1301	GGAAACAGTGGCTCGCGGAAGACGACCGTGCAGCTGATGCAAGAGGCTCTATGACCCC	1360
QY	379	AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr	398
DB	1361	ACAGATGGCATGTGCTGTAATGATGACAGGACATTAGCACCAATAATGTAAGGCATCTT	1420
QY	399	ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn	418
DB	1421	CGGAAATTAATCTGCTGTGTGAGTCAGGAGCCTGTGTTGTTGTCACCCAGCATGTAA	1480
QY	419	AsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAaArgGlu	438
DB	1481	AACATTCGCTATGCGCGGCAAAATGTCCACATGAGTGAAGTGTGAAGAGCTGTTAAAGNA	1540
QY	439	AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu	458
DB	1541	GCCATGCCCTATGATTTTATGAAACTTACCTAATAATTTGACACTCTGCTGGTTGGAGAG	1600
QY	459	LysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal	478
DB	1601	AGAGGGGCCCCAGCTGAGTGTGGACAGAAACAGAGAATCGCCATTGCTCGGGCCCTGGT	1660
QY	479	ArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys	498
DB	1661	CGCAACCCCAAGATTCTCTGCTGGATGAGGCAACGCTCAGCTCTGGACACTCGAAAGTGAA	1720
QY	499	SerAlaValGlnAlaAlaLeuGlyLysAlaSerLysGlyArgThrThrIleValValAla	518
DB	1721	CGATGTTTTCAGTGGCCCTGGTAAGGCCAGAAAGGCCGAGCTACCATTTGTATAGCT	1780
QY	519	HisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeu	538
DB	1781	CATCGTTTGTCTACAGTTCGTAATGCCGATGTCATTTGCTGGTTTGATGATGGAGTCATT	1840
QY	539	AlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuVal	558
DB	1841	GTGGAGAAAGGAAATCATGATGAACATCATGAAAGAGAGAGGGCATTTTACTTCAAACTTGT	1900
QY	559	MetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu	578
DB	1901	ACAATGTCAG---ACAAGAGAAATGAAATTTGAGTTAGAAATGCCACCTGGTGAATCCAAA	1957
QY	579	ArgLysThrAsnSerLeuHisSerValLysSerIleLysSerAspPheIle---	597
DB	1958	AGTGAAGTGTATGCTTGGAAATG---TCTCCAAAGAGATTGAGGCTCCAGTTTAATAAAA	2014
QY	598	-----AspLysAlaGluGluSer	603
DB	2015	AGAAGATCAACTCGCAGGAGTATACATGCACCACCAAGGGCCAAAGACAGAAAGCTTGGTACA	2074

QY	604	ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn	623
DB	2075	AAAGAGGACTTGAATGAGAAATCACTCCAGTTTTCTCTGAGGAAATCTCGAAGCTGAAC	2134
QY	624	LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal	643
DB	2135	TCAACTGAATGGCCCTATTTTGGTGGTATATTTTGGCTATATATAACGAGGCCCTG	2194
QY	644	HisProValPheSerIleIlePheAlaLysIleIleThrMetPhe--GlyAsnAsnAsp	662
DB	2195	CAACACGAGCATTTTCAATAATATTTTCAGGATATATAGGGATCTTTACCCGAGATGAGAT	2254
QY	663	LysThrThrLeuLysHisAspAlaGluIleThrSerMetIlePheValIleLeuGlyVal	682
DB	2255	CCTGAACAAAAACGACAGAAATAGTAACATGTTTCTGTATTTGTTCTAGTCCTTGAAT	2314
QY	683	IleCysPheValSerThrPheMetGlnGlyLeuPheThrGlyArgAlaGlyGluIleLeu	702
DB	2315	ATTCTCTTTATTACATTTTCTCCAGGCTTCACATTTGGCAACGCTGGGAGATCCTC	2374
QY	703	ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuThrGlnAspIleAlaTrpPhe	722
DB	2375	ACTAAGCGCTTCGATACATGTTTTCAGATCCATCTCGACACGAGATGTCAGCTGGTTT	2434
QY	723	AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln	742
DB	2435	GATGACCCCTAAAAACACCACTGGAGCATTTGACAAACAGCGCTGGCCAAATGATCGGCTCAA	2494
QY	743	IleGlnGlyAlaThrClySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly	762
DB	2495	GTTAAAGGGGCTATAGTTCCAGGCTTGCTGTCATTACCACGAAATATAGCAATCTTGGG	2554
QY	763	LeuSerValIleIleSerPheIleThrGlyTrpGluMetThrPheLeuIleLeuSerIle	782
DB	2555	ACAGGCAATTATATATCCTTATCTATGTTGGCAATTAACACTTTTACTCTTAGCAATT	2614
QY	783	AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn	802
DB	2615	GTACCCATCATTTGCAATAGCAGGAGTGTGTGAATGAAATGTTGTGTGGACAGCACTG	2674
QY	803	LysAspLysGlnLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle	822
DB	2675	AAAGATTAAGAAAGAGCTAAGAGAGCTGGGAAGATTGCTACAGAGCCATCGAANAATTC	2734
QY	823	ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetThrGluGluMetLeu	842
DB	2735	CGNACTGTTGTTCTTTGACTCGGAGCAGAGAGTTTGAATACATGATGCACAGAGTTTG	2794
QY	843	GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysThrAlaPhe	862
DB	2795	CAAGTACCATACAGAAACTCTTTGAGGAAGACACACATCTTCGGGGTCTCATTTCTTATC	2854
QY	863	SerHisAlaPheIleThrPheAlaThrAlaAlaGlyPheArgPheGlyAlaThrLeuIle	882
DB	2855	ACCAGCGCAATGATGATTTTCTTATGCTGGCTGTTTTCGGTTGGTGCTACTTCGGTG	2914
QY	883	GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaThrGly	902
DB	2915	GCMAATGAGTTCAAGAACTTTCCAGGATGTTCTTTGGTATCTTCAGCTATGCTTTGGT	2974
QY	903	AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluThrSerLysAlaLysSerGly	922
DB	2975	GCCATGGCAATGGGCAGGTCAGTTTCATTTGCTCCTGACTATGCAAGGCCAAGATATCA	3034
QY	923	AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu	942
DB	3035	GCAGCCACGCTCATCATGATCATTTGAAAGGCCCTCTGATTGACAGCTACAGCCCTCAC	3094
QY	943	GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheThr	962
DB	3095	GGCCTCAGGCCAATACGTTGGAGGAAATGTGACATTTAATGAGTCTGTTCAACTAT	3154
QY	963	ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys	982


```
Db 3155 CCACCTGACACAGACATCCCGTGTCTCAGGGGCTGAGCTCGAGTGAAGAGGCCAG 3214
Qy 983 ThrValAlaPheValGlySerGlyCysGlyLysSerThrSerValGlnLeuGln 1002
Db 3215 ACGTGGCCCTCGTAGTAGAGTGGCTGTGGGAAGACACAGTGTTCAGCTCTAGAG 3274
Qy 1003 ArgLeuTyrAspProValGlnGlnValLeuPheAspGlyValAspAlaLysGluLeu 1022
Db 3275 CGCTTCTATGACCCCTGTGCTGTTCAGTCTAAATTGATGGCAAGAGATAAAGCACCTG 3334
Qy 1023 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 1042
Db 3335 AATGTCCAGTGGCTCGAGACACCTGGGCATCGTCTCAGAGGCCATCTGTTTGAC 3394
Qy 1043 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 1062
Db 3395 TGCAGATTGCCGAGAACATTGCCATATGGAGACACAGCCGGTCTGATCACATGAAGAG 3454
Qy 1063 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 1082
Db 3455 ATTATGCGAGCGACCAAGAGGCCAACATACACCACTTCATCGACACACTCCCTGAGAA 3514
Qy 1083 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 1102
Db 3515 TACAACACAGAGTAGGACAAAGAACCCAGCTCTCTGGTGGCCAGAAACAGCGCAT 3574
Qy 1103 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 1122
Db 3575 GCCATAGCTCGGCTCTGTGTAGACAGCTCATATTTCTTTGGATGAAGTACATCA 3634
Qy 1123 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 1142
Db 3635 GCTCTGGATACAAAGTGAAGGTTGTCCAAGAGCCCTGGACAAAGCCAGAGAAGGC 3694
Qy 1143 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 1162
Db 3695 CGCAGCTGCAATGTGATGCCCGCCAGCTGTCCACCATCCAGATGCAGATTAAATAGTG 3754
Qy 1163 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 1182
Db 3755 GTGTTTCAGATGGCAAGTCAAGGAGCATGGCACACATCAACAGCTGTGCCCCAGAA 3814
Qy 1183 AspIleTyrPheLysLeuValAsnAlaGlnSer 1193
Db 3815 GGCATCTATTTTCCATGCTCAGTGTCCAGGCT 3847

RESULT 7
AAD03505
ID AAD03505 standard; cDNA; 4279 BP.
XX
AC AAD03505;
XX
DT 13-JUN-2001 (first entry)
XX
DE Dog P-glycoprotein (PGP) allelic variant (Genotype B) cDNA.
XX
KW Dog; P-glycoprotein allelic variant; PGP; multidrug transporter;
MDRI; drug bioavailability; transgenic animal; genetic model; ss.
XX
OS Canis familiaris.
XX
FH Key Location/Qualifiers
CDS 17..3862
FT /*tag= a
FT /product= "Dog P-glycoprotein (PGP) allelic variant
FT (Genotype B) protein"
FT replace (91, T)
FT /*tag= b
FT replace (607, C)
FT /*tag= c
XX
PN WO200123540-A2.
```

```
XX 05-APR-2001.
XX 28-SEP-2000; 2000WO-US26767.
XX 28-SEP-1999; 99US-0156510.
XX (GENT-) GENTEST CORP.
XX Stocker PU, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;
XX WPI; 2001-235373/24.
XX P-PSDB; AAE00309.
XX New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
PT for determining the bioavailability of drugs and for screening for dog
PT PGP inhibitors -
XX Claim 9; Page 93-99; 11lpp; English.
XX The invention relates to dog P-glycoprotein (PGP) also referred
CC as multidrug transporter (MDRI) and nucleic acids encoding them.
CC The invention also includes fragments and biologically functional
CC variants of dog P-glycoprotein. PGP and their nucleic acids are
CC useful for determining the bioavailability of drugs and for
CC screening PGP inhibitors. They are useful for the diagnosis and
CC treatment of conditions characterised by PGP activity, by
CC reducing or increasing PGP activity in a cell. PGP nucleic acids
CC are used as oligonucleotide probes. Complements of PGP nucleic
CC acids are useful as antisense oligonucleotides, to induce a PGP
CC 'knockout' phenotype. They are used to prepare a non-human
CC transgenic animal, which are valuable as genetic models for
CC human diseases.
CC The present sequence is dog P-glycoprotein (PGP) allelic variant
CC (Genotype B) cDNA. The PGP enzyme functions as an efflux pump
CC exporting small molecules across the cell membrane. This enzyme
CC is a member of the ABC transporter family.
XX Sequence 4279 BP; 1296 A; 833 C; 1008 G; 1142 T; 0 other;

Alignment Scores:
Pred. No.: 0 Length: 4279
Score: 3554.00 Matches: 682
Percent Similarity: 74.74% Conservative: 238
Best Local Similarity: 55.40% Mismatches: 269
Query Match: 59.12% Indels: 42
DB: 22 Gaps: 8

US-09-873-409-6 (1-1195) x AAD03505 (1-4279)
Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 167 ATGTTGGTGGGAGCAATGCTGCCATCATCCATGAGCTGCACTCCCTCTCATGATGCTG 226
Qy 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu----- 34
Db 227 GTTTTGGAAACATGACAGATAGCTTTGCAATGCGAGGAATTTCAAGAAACAAACTTTT 286
Qy 35 -----ValGlnThrAsnThrTyrSerPhePheArg----- 44
Db 287 CCAGTTATAATTAATGAAGTATTACGAAACATACACAACTTCATCAACCATCTGGAG 346
Qy 45 -----LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe 59
Db 347 GAGGAATGACCCACGATGCTCTATTATTACAGTGGGATCGGTGCTGGCTGGCT 406
Qy 60 GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg 79
Db 407 GCTTACATCCAGTTTTCATTCTGTCCTGCGCAGCAGGAGACAGATACTCAAAATTAGA 466
Qy 80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99
Db 467 AAACAATTTTTCATGCTATCATGCGAGGAGATTGGCTGTTTGACGTGCATGACGTT 526
```


Qy	100	GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp	118
Db	527	GGGAGCTTAACACCCGGCTCAGACGATGCTCCAAAATCAATGAAGAAATTGGCGAC	596
Qy	119	LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu	138
Db	587	AAAATTGGAATTCTTTTCAATCAATAGCAACATTTTTCACCGGTTTTATAGTGGGTTT	646
Qy	139	VallysglyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer	158
Db	647	ACACGTGGTTGGAAAGCTAACCCCTTGATTTTGGCCATCAGCCCTGTTCTTGGACATTCA	706
Qy	159	AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer	178
Db	707	GCGCCATCTGGCAAGATACATCTTCACTTACTGATAAAGAACTCTTGGCCCTATGCA	766
Qy	179	LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg	198
Db	767	AAAGCTGGACAGTAGCTGAAGAAGTCTTAGCAACAATCAGAACTGTGATTCCTTTGGGA	826
Qy	199	AlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly	218
Db	827	GGACAAAGAAAGAACTTGAAGAGTACACAAAATTTAGAAAGACTTAAGAGTAATTTGGG	886
Qy	219	IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGly	238
Db	887	ATAAGAAGACTATCAGGCCAACATTTCTATTTGGTCGCGCTTCTTATGATCATGTCGA	946
Qy	239	ThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr	258
Db	947	TCATATGCTCTGGCTTCTGGTATGGACCTCCTTGGTCTCTCCAGTGAA-----TAT	1000
Qy	259	ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly	278
Db	1001	TCTATTGGCAAGTACTCACTGTCTCTTTTCTGTATTAAATTGGGGCTTTTAGTATTGGA	1060
Qy	279	AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe	298
Db	1061	CAGCATCCCCAAGCATTTGAAGCATTTTGCAACGACAGAGAGAGCTTATGAATCTTC	1120
Qy	299	GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu	318
Db	1121	AAGATAATTGCAATAAACCAACATTTGACAGCTATTGCAAGAGTGGACATAAACACGAT	1180
Qy	319	SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer	338
Db	1181	AATATTAAAGGAAATTTGGAAATTCAAAAATGTTTCACTTCAGTTACCCCTCTCGAAAAA	1240
Qy	339	IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal	358
Db	1241	GTTAAGATCTTTAAGGCTCTCAACCTGAAGGTTTCAGATGGCGACAGTAGTGGCCCTGGT	1300
Qy	359	GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro	378
Db	1301	GGGAACAGTGGCTCGGGAAGACACGACCGTGCAGCTGATGCAGAGGCTCTATGACCCC	1360
Qy	379	AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr	398
Db	1361	ACAGATGGCATGTGCTGTATTGATGGACAGGACATTTAGGACCATAAATGTAAGGCATCTT	1420
Qy	399	ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn	418
Db	1421	CGGGAATTTACTGTGTGGTGGTCTCAGGAGCCTGTGTTTGGCCACCACCATGATGTAA	1480
Qy	419	AsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGlu	438
Db	1481	AACATTGCTATGGCCGCGAAAATGTGCACCATGGATGAGATTGAGAAAGCTCTTAAGGAA	1540
Qy	439	AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu	458
Db	1541	GCCAAATGCATGATTTTATCATGAAATACCTTAATAAATTTGACATCTCGTTGGGAGAG	1600

Qy	459	LysGlyAlaGlnMetSerGlyGlyGlnIysGlnArgIleAlaIleAlaArgAlaLeuVal	478
Db	1601	AGAGGGGCCAGCTAGTGGTCGACAGAACAGAGAAATCGCCATTGCTCGGGCCCTGGT	1660
Qy	479	ArgAsnProIysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerIys	498
Db	1661	CGCAACCCCAAGATTCTTCGTGGATGAGGCAACGTGAGCTCTGGACACTGNAAGTGAA	1720
Qy	499	SerAlaValGlnAlaLeuGluIysAlaSerIysGlyArgThrThrIleValValAla	518
Db	1721	GCAGTGGTTCAGGTGGCCCTGGNTAAGCCAGAAAGCCGGGACTACCAATGTGATAGCT	1780
Qy	519	HisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuIysAspGlyMetLeu	538
Db	1781	CATCGTTGTCTACAGTTCGTAAATGCCGATGTCAATTCGTGGTTTGTATGATGAGTCAAT	1840
Qy	539	AlaGluIysGlyAlaHisAlaGluLeuMetAlaIysArgGlyLeuTyrTyrSerLeuVal	558
Db	1841	GTGGAGAAAGGAATCATGATGAACTCATGAAGAGAGAGGGCATTTACTTCAAACTTGTC	1900
Qy	559	MetSerGlnAspIleIysAlaAspGluGlnMetGluSerMetTyrSerThrGlu	578
Db	1901	ACAATGCAG---ACAAGAGGAATGAAATTGAGTTAGAAATGCCACTGGTGAATCCAA	1957
Qy	579	ArgIysThrAsnSerLeuProLeuHisSerValIysSerIleIysSerAspPheIle---	597
Db	1958	AGTGAAGTGAATGCTTGGAAATG---TCTCCAAAGATTCAGGGTCCAGTTTAATAAAA	2014
Qy	598	-----AspLysAlaGluIleSer	603
Db	2015	AGAAATCAACTCGCAGGAGTATACATGCACCACAGGCCAAGACAGAAAGCTGGTACA	2074
Qy	604	ThrGlnSerIysGluIleSerLeuProGluValSerLeuLeuIysIleLeuIysLeuAsn	623
Db	2075	AAAGAGACTTGAATGAGAAATGATCACTCAGTTTCTTCTGGAGGATTCAGACTGAAC	2134
Qy	624	LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal	643
Db	2135	TCAACTGAATGGCCTTATTTGTGGTTGGTATATATTTGTGCTATTATAACCGAGGCCCTG	2194
Qy	644	HisProValPheSerIleIlePheAlaIysIleIleThrMetPhe---GlyAsnAsnAsp	662
Db	2195	CAACAGCAATTTTCAATAATATTTTCAAGGATATATAGGATCTTTACCAGAGTAGAGAT	2254
Qy	663	LysThrThrLeuIysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal	682
Db	2255	CCTGAACAAACACAGAAATAGTAACTGTTTTCTGTATGTTTCTAGTCCCTGGGAAT	2314
Qy	683	IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyIleLeu	702
Db	2315	ATTTCTTTTATACATTTTCTCCAGGCTTCACATTTGGCAAAGCTGGGAGATCCTC	2374
Qy	703	ThrMetArgLeuArgHisLeuAlaPheIysAlaMetLeuTyrGlnAspIleAlaTrpPhe	722
Db	2375	ACTAAGCGGCTTCGATACATGGTTTTTTCAGATCCATGCTGAGACAGGATGTCAAGCTGG	2434
Qy	723	AspGluIysGlnAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln	742
Db	2435	GATGACCTTAACAAACCACTGGAGCATTCACACCAGGCTTGCAATGATGCGGCTCA	2494
Qy	743	IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly	762
Db	2495	GTAAAGGGGCTATAGTTCAGGCTTGCTGTCATTTACCAGAAATATAGCAAAATCTTGGG	2554
Qy	763	LeuSerValIleIleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIle	782
Db	2555	ACAGCAATATATATCTTAAATCTATGTTGGCAATTAACACTTTTACTCTTACGAAT	2614
Qy	783	AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn	802
Db	2615	GTACCCATCATGCAATAGCAGGAGTGTGTAAATGAATAATGTTGCTGGACAGACACTG	2674
Qy	803	LysAspIysGlnGluLeuIysHisAlaGlyIysIleAlaThrGluAlaLeuGluAsnIle	822

[illegible]

D	b	3755	GTGTTTCAGAAATGGCAAGTCAAGGAGCATGGCACACATCAACAGCTGCGCCAGAAA	3814
Q	y	1183	AspiletyrPhelysLeuValAenLaGlnSer	1193
			() : :: :::	
D	b	3915	GGCATTATTTTCCATGGTGTCAGTGCCAGGCT	3947
R	E	S	U	L
A	A	A	A	A
I	D	AAD03488	standard; cDNA; 4279 BP.	
X	X	AC	AA03488;	
X	X	DT	13-JUN-2001 (first entry)	
X	X	DE	Dog P-glycoprotein (PGP) cDNA #1.	
K	W	Dog; P-glycoprotein; PGP; multidrug transporter; MDR1;		
K	W	drug bioavailability; transgenic animal; genetic model; ss.		
X	X	OS	Canis familiaris.	
X	X	FH	Key Location/Qualifiers	
F	T	CDS	17..3862	
F	T		/tag= a	
F	T		/product= "Dog P-glycoprotein (PGP) #1"	
X	X	FN	WO200123540-A2.	
X	X	PD	05-APR-2001.	
X	X	PF	28-SEP-2000; 2000WO-US26767.	
X	X	PR	28-SEP-1999; 99US-0156510.	
X	X	PA	(GENT-) GENTEST CORP.	
P	I	Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;		
D	R	WPI; 2001-235373/24.		
D	R	P-PSDB; AAE0303.		
F	T	New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful		
P	T	for determining the bioavailability of drugs and for screening for dog		
X	X	PGP inhibitors -		
P	S	Claim 3; Page 58-63; 11pp; English.		
C	C	The invention relates to dog P-glycoprotein (PGP) also referred		
C	C	as multidrug transporter (MDR1) and nucleic acids encoding them.		
C	C	The invention also includes fragments and biologically functional		
C	C	variants of dog P-glycoprotein. PGP and their nucleic acids are		
C	C	useful for determining the bioavailability of drugs and for		
C	C	screening PGP inhibitors. They are useful for the diagnosis and		
C	C	treatment of conditions characterised by PGP activity, by		
C	C	reducing or increasing PGP activity in a cell. PGP nucleic acids		
C	C	are used as oligonucleotide probes. Complements of PGP nucleic		
C	C	acids are useful as antisense oligonucleotides, to induce a PGP		
C	C	'knockout' phenotype. They are used to prepare a non-human		
C	C	transgenic animal, which are valuable as genetic models for		
C	C	human diseases.		
C	C	The present sequence is dog P-glycoprotein (PGP) cDNA. This		
C	C	sequence is also referred as Genotype C cDNA. The		
C	C	PGP enzyme functions as an efflux pump exporting small molecules		
C	C	across the cell membrane. This enzyme is a member of the ABC		
C	C	transporter family.		
S	Q	Sequence 4279 BP; 1294 A; 834 C; 1008 G; 1143 T; 0 other;		
A	l	i	g	n
A	l	i	g	n
S	c	r	e	:
P	r	e	d. No.:	0
S	c	r	e:	Length: 4279
P	e	r	cent Similarity:	3549.00
B	e	s	t Local Similarity:	74.65%
M	i	s	matches:	Matches: 681
M	i	s	mismatches:	Conservative: 238
M	i	s	mismatches:	Mismatches: 270

Alignment Scores:		
Pred. No.:	0	Length: 4279
Score:	3549.00	Matches: 681
Percent Similarity:	74.65%	Conservative: 238
Best Local Similarity:	55.32%	Mismatches: 270

Query Match: 59.03% Indels: 42
DB: 22 Gaps: 8
US-09-873-409-6 (1-1195) x AAD03488 (1-4279)

QY 1 MetIleuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
DB 167 ATGTGGTGGGCAATGGCTGGCATCATCGATGGAGCTCACTCCCTCTCATGATGCTG 226
QY 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu----- 34
DB 227 GTTTTGGAAACATGACAGATAGCTTTGCAAAATGCGAGAAATTCAGAAACAAAACTTTT 286
QY 35 -----ValGlnThrAsnThrThrThrThrThrPheArg----- 44
DB 287 CCAGTTAATAATGAAGATTAATACGAACAATACACAATTCATCAACCATCTGGAG 346
QY 45 -----LeuThrLeuTyrTyrValGlyIleGlyValAlaLeuIlePhe 59
DB 347 GAGGAATGACCACTATGCTTATTTACAGTGGATCGGTGCTGGCTGCTGCTGCT 406
QY 60 GlyTyrIleGlnIleSerLeuTyrIleThrAlaAlaArgGlnThrLysArgIleArg 79
DB 407 GCTTACATCCAGGTTTCATTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 466
QY 80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIle 99
DB 467 AACAAATTTTTCATGCTATCATCGACAGAGATTTGGCTGGTTTGACGGTCATGAGTT 526
QY 100 GlycLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
DB 527 GGGGAGCTTAACACCCGGCTCACAGACGATGCTCCAAAATCAATGAAGAAATGGCGAC 586
QY 119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
DB 587 AAAATTGGAATGTTCTTCTCACTCAATAGCAACATTTTTCACCGGTTTATAGTGGGGTTT 646
QY 139 ValLysGlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
DB 647 ACAGTGGTTGGAAGCTAACCCCTTGATTTGGCCATCAGCCCTGCTCTGGACITTC 706
QY 159 AlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
DB 707 GCGCCATCTGGGCAAGATACTATCTTCACTTACTGATAAAGAACTCTTGGCCTATGCA 766
QY 179 LysAlaGlyAlaValAlaGluGlnValLeuSerSerIleArgThrValIleAlaPheArg 198
DB 767 AAAGCTGGAGCAGTAGCTGAAGAAGCTTTAGCAGCAATCAGAACTGTGATTCCTTTTGA 826
QY 199 AlaGlnGlyLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 218
DB 827 GGACAAAGAAAGAACTTGAAGGTACACAAATAATTTAGNAGAGTAAGGAAATTTGGG 886
QY 219 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPheMetAsnGly 238
DB 887 ATAAAGAAAGCTATCACGGCCAACATTTCTATTGTCGGCTTCTTATTGATCATGCA 946
QY 239 ThrTyrGlyLeuAlaPheThrTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 258
DB 947 TCATATGCTCTGGCTTCTGATGGAGACCTCTTGTGCTCTCTCCAGTGAA-----TAT 1000
QY 259 ThrIleGlyThrValLeuAlaValPheSerValIleHisSerSerTyrCysIleGly 278
DB 1001 TCTATTGGACNAGTACTACATGCTCTTCTTCTGATTAATTTGGGGCTTTTAGTATTGGA 1060
QY 279 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 298
DB 1061 CAGGCATCCCAAGCACTTGAAGCATTTTGCAACGCAAGAGAGAGCAGCTTATGAAATCTTC 1120
QY 299 GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 318
DB 1121 AAGATAATTGACATAACCAAGCAATTTGACAGCTATTTCGAAGAGTGGACATAAACCCAGAT 1180

QY 319 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 338
DB 1181 AATATTAGGGAAATTTGGAAATTCAAAAATGTTCACTTCAGTTACCTTTCTCGAAAGAA 1240
QY 339 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 358
DB 1241 GTTAAGATCTTAAGGGTCTCAACCTGAAGGTTCAAGTGGGAGAGTGGCGCTGGTT 1300
QY 359 GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro 378
DB 1301 GGGAAACAGTGGCTGGGGAAGAGACACCGTGCAGCTCATGCAGAGGCTCTATGACCCC 1360
QY 379 AspAspGlyPheIleMetValValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 398
DB 1361 ACAGATGGCATGCTGTATTTGATGACAGACATTTAGACCAATAATGTAAGGCATCTT 1420
QY 399 ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn 418
DB 1421 CGGAAATTTACTGGTGTGTGAGTCAGGAGCCTGTGTTTGGCCACCATGATGCTGAA 1480
QY 419 AsnIleLysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGlu 438
DB 1481 AACATTCGTATGGCGCGCAAAATGTCACCATGGATGAGATTGAGAAGCTGTTAAGAA 1540
QY 439 AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 458
DB 1541 GCCAATGCCTATGATTTTATCATGAACCTACCTAATAAATTTGACACTCTCTGGTTGGAG 1600
QY 459 LysGlyAlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal 478
DB 1601 AGAGGGCCCGCTGAGTGGTGGACAGAAACAGAGAAATCGCCATGCTCGGGCCCTGGTT 1660
QY 479 ArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 498
DB 1661 CGCAACCCCAAGATCTTCTGCTGGATGAGCAACCTGAGCTCTGGACACTGAAAGTAA 1720
QY 499 SerAlaValGlnAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAla 518
DB 1721 GCAGTGGTTGAGTGGCTGGATAGGCGCAGAAAGGCGGAGTACCATCTGTGTAGTAGCT 1780
QY 519 HisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeu 538
DB 1781 CATCGTTGTGTACAGTTCGTAAATGCCGATGTCATTCGTGTTTTCATGATGAGTCAIT 1840
QY 539 AlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuVal 558
DB 1841 GTGGAGAAAGAAATCATGATGNACTCATGAAGAGAGAGGCACTTACTTCAAACTGTC 1900
QY 559 MetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu 578
DB 1901 ACAATGCAG---ACAAGAGGAAATGAATAGTTAGTAAATAATGCCACTGTGTAATCCAA 1957
QY 579 ArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle--- 597
DB 1958 AGTGAAGTGTATGCTTGGAAATG---TCTCCAAAAGATTACGGGTCAGTTTAATAAAA 2014
QY 598 -----AspLysAlaGluGluSer 603
DB 2015 AGAAGATCAACTCGCAGGAGTATACATGCACCACACAGGCCAACAGCAAGCTGGTACA 2074
QY 604 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLysIleLeuLysLeuAsn 623
DB 2075 AAAGAGGACTTGAATGAGAATGTACCTCCAGTTTCTCTTGGAGGATTTCTGAAGCTGAAC 2134
QY 624 LysProGluTyrProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 643
DB 2135 TCAACTGAATGGCTTATTTTGGTGGTATATTTTGTGCTATTATATAACGAGGCGCTG 2194
QY 644 HisProValPheSerIleIlePheAlaLysIleIleThrMetPhe---GlyAsnAsp 662
DB 2195 CAACCGACATTTTCAATATATATTTTCAAGGATATAGGATCTTTACCCGAGATGAGAT 2254
QY 663 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 682


```
Db 2255 CCTGAACAAAACGACAGAAATAGTAACATGTTTCTGTATTGTTCTAGTCTCTTGAATT 2314
Qy 683 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluLeuLeu 702
Db 2315 ATTTCTTTTATTACATTTTCTCCAGGGCTTCACATTTGGCAAGCTGGAGATCTCT 2374
Qy 703 ThrMetArgLeuArgHisLeuAlaPheTyrAlaMetLeuTyrGlnAspIleAlaTrpPhe 722
Db 2375 ACTAAGCGGCTTCGATACATATGTTTTCAGATCCATCTGAGACAGGATGTCAGCTGGTT 2434
Qy 723 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 742
Db 2435 GATGACCCCTAAAAACACCACTGGAGCATTTGCACAAACAGGCTTGCCCAATGATCGGCTCAA 2494
Qy 743 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 762
Db 2495 GTTAAAGGGGCTATAGTTCCAGGCTTGCTGTCATTTACCCAGATATAGCAATCTTGGG 2554
Qy 763 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle 782
Db 2555 ACAGGCAATTATTATATCTTAATCTATCTGTTGGCAATTAACACTTTTACTCTTAGCAATT 2614
Qy 783 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 802
Db 2615 GTACCCATCATTTGCAATAGCAGAGTTGTTGAATGAAATGTTGTCGCAACAGCACTG 2674
Qy 803 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGlnAlaLeuGluAsnIle 822
Db 2675 AAGATPAGAAGAGCTAGAGAGCTGGGAAGTTGCTACAGAACCACTCGAAACTTC 2734
Qy 823 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 842
Db 2735 CGAAGCTGTTGTTCTTTGACTCGGAGCAGAGTTTGAATACATGATGATCACAGAGTTTG 2794
Qy 843 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 862
Db 2795 CAAGTACCATACAGAAACTCTTTAGAGAAAGCACACATCTTCGGGGTCTCAATTTCTATC 2854
Qy 863 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 882
Db 2855 ACCAGGCAATGATGATTTTCTCATGTGCTGCTGTTTCCGGTTTGGTCTACTTGGTG 2914
Qy 883 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 902
Db 2915 GCAAATGAGTTTCATGAACTTCAGGATGTTCTTTTGGTATCTCAGCTATTGCTTTGGT 2974
Qy 903 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 922
Db 2975 GCCATGGCAGTGGGCGAGGTCACTTCAATTTGCTGCTGACTATGCCAAAGCCAAAGTATCA 3034
Qy 923 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 942
Db 3035 GCAGCCAGCTCATCATGATCATTTGAAAGAAAGCCCTCTGATTGACAGCTACAGCCCTCAC 3094
Qy 943 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPheTyr 962
Db 3095 GGCCTCAGCCAAATACGTTGGAGGAATGTGACATTTAATGAGTCCGTTTCACTAT 3154
Qy 963 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 982
Db 3155 CCCACTCGACACACATCCCGTCTCCAGGGCTGAGCCTCGAGGTGAAGAAGGGCCAG 3214
Qy 983 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 1002
Db 3215 ACCTGCGCCCTCGTAGTACAGTGGCTGGGAAGAGACAGCTTGTGAGTCTCCCTAGAG 3274
Qy 1003 ArgLeuTyrAspProValGlnGlnValLeuPheAspGlyValAspAlaLysGluLeu 1022
Db 3275 CGTTCTATGACCCCTTGGCTGTTGCTGCTAATGATGGCAAGAGATAAGACACCTG 3334
Qy 1023 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 1042
```

```
Db 3335 AATGTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCTCAGAGGCCCATCTGTTTGAC 3394
Qy 1043 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 1062
Db 3395 TGCAGCATTTGCCGAGAACATTGCTTATGAGACAAACAGCCGGTCTGATCATCATGAAGAG 3454
Qy 1063 IleLysGluAlaAlaAsnAlaIleAsnIleHisSerPheIleGluGlyLeuProGluLys 1082
Db 3455 ATTATGCGAGGCCCAAGGAGGCCAACATACACCTTTCATCGAGACACTCCCTTGAGAAA 3514
Qy 1083 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 1102
Db 3515 TACAACACCGAGTAGGACACAAGAACCCAGCTCTCTGTGGCCAGAAACAGCGCAT 3574
Qy 1103 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuAspGluAlaThrSer 1122
Db 3575 GCATAGCTCGGCTCTTGTGTAGACAGCCTCATATTTTGTGATGAAGCTACATCA 3634
Qy 1123 AlaLeuAspAsnAspSerGlnLysValGlnHisAlaLeuAspLysAlaArgThrGly 1142
Db 3635 GCTCTGGATACAGAAGTGAAGGTTGTCCAAGAGGCCCTCGACAAAGCCAGAGAGGC 3694
Qy 1143 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 1162
Db 3695 CGCACTGCTATGTATGCCCCACCGCTTGTCCACATCCAGATGCAGATTTAATAGTG 3754
Qy 1163 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 1182
Db 3755 GTGTTTCAGAAATGGCAAGTCAAGGAGCATGCGACACATCAACAGCTGTGCGCCAGAAA 3814
Qy 1183 AspIleTyrPheLysLeuValAsnAlaGlnSer 1193
Db 3815 GGCATCTATTTTCCATGTGTCAGTGTCCAGGCT 3847

RESULT 9
AAD03489
ID AAD03489 standard; cDNA; 4317 BP.
XX
AC AAD03489;
XX
DT 13-JUN-2001 (first entry)
XX
DE Dog P-glycoprotein (PGP) cDNA #2.
XX
KW Dog; P-glycoprotein; PGP; multidrug transporter; MDR1;
KW drug bioavailability; transgenic animal; genetic model; ss.
XX
OS Canis familiaris.
XX
FH Key Location/Qualifiers
FT CDS 70..3912
FT /*tag= a
FT /product= "Dog P-glycoprotein (PGP) #2"
XX
PN WO200123540-A2.
XX
XX 05-APR-2001.
XX
XX 28-SEP-2000; 2000WO-US26767.
XX
XX 28-SEP-1999; 99US-0156510.
XX
XX (GENT-) GENTEST CORP.
XX
XX Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;
XX WPI; 2001-235373/24.
XX P-PSDB; AAE00304.
XX
XX New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
XX for determining the bioavailability of drugs and for screening for dog
XX PGP inhibitors -
XX
```


Claim 1; Page 66-72; 11lpp; English.

XX The invention relates to dog P-glycoprotein (PGP) also referred
 CC as multidrug transporter (MDR1) and nucleic acids encoding them.
 CC The invention also includes fragments and biologically functional
 CC variants of dog P-glycoprotein. PGP and their nucleic acids are
 CC useful for determining the bioavailability of drugs and for
 CC screening PGP inhibitors. They are useful for the diagnosis and
 CC treatment of conditions characterized by PGP activity, by
 CC reducing or increasing PGP activity in a cell. PGP nucleic acids
 CC are used as oligonucleotide probes. Complements of PGP nucleic
 CC acids are useful as antisense oligonucleotides, to induce a PGP
 CC 'knockout' phenotype. They are used to prepare a non-human
 CC transgenic animal, which are valuable as genetic models for
 CC human diseases.

CC The present sequence is dog P-glycoprotein (PGP) cDNA. The
 CC PGP enzyme functions as an efflux pump exporting small molecules
 CC across the cell membrane. This enzyme is a member of the ABC
 CC transporter family.

XX Sequence 4317 BP; 1293 A; 844 C; 1019 G; 1161 T; 0 other;

Alignment Scores:

Pred. No.:	0	Length:	4317
Score:	3549.00	Matches:	680
Percent Similarity:	74.65%	Conservative:	239
Best Local Similarity:	55.24%	Mismatches:	270
Query Match:	59.03%	Indels:	42
DB:	22	Gaps:	8

US-09-873-409-6 (1-1195) x AAD03489 (1-4317)

QY	1	MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu	20
DB	217	ATGTGGTGGGACATAGGGCTGCATCATCCATGAGCTGCACCTCCTCATGATGCTG	276
QY	21	ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu	34
DB	277	GTITTTGGAACATGACAGATAGCTTTGCAATGCAGGAATTTCAAGAAACAAACTTTT	336
QY	35	-----ValGlnThrAsnThrTyrSerPhePheArg-----	44
DB	337	CCAGTTATAATTAATGAAGTAATTAACCAACATACACACATTTTCACACCATCTGGAG	396
QY	45	-----LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuPhe	59
DB	397	GAGAAATACCCAGTATGCTATTAACAGTGGATCGGCTGCGCTGCTGGTGGCT	456
QY	60	GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg	79
DB	457	GCTTACATCCAGGTTTCATCTGCTGCTGGCAGCAGGAAGACAGATACTCAAAATTAGA	516
QY	80	LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle	99
DB	517	AAACAATTTTTCATGCTATCATCGCACAGAGATTGGCTGGTTGACGTGCATGACGTT	576
QY	100	GlyGluLeuAsnThrArgMetThr--AspIleAspIleSerAspGlyIleGlyAsp	118
DB	577	GGGAGCTTAACACCCGGCTCACAGCATGTCTCCAAAATCAATGAAGAAATGGCGAC	636
QY	119	LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu	138
DB	637	AAAGTTGGAATGTTCTTCAATCAATAGCAACATTTTTCACCGGTTTATAGTGGGTTT	696
QY	139	ValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer	158
DB	697	ACACCTGGTTGGAAGCTTAACCCCTGTGATTTTGGCCATCAGCCCTGTTCTTGGCACTTCA	756
QY	159	AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer	178
DB	757	GCCGCCATCTGGGCAAGATCATCTTCACTTACTGATGAAGAACTCTTGGCCATGCA	816
QY	179	LysAlaGlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArg	198

DB	817	AAAGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGAACTGTGATTGCTTTGGA	876
QY	199	AlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly	218
DB	877	GGACAAAGAAAGAACTTGAAGGTACAAACAAATTTAGAAAGACTAAAGAAATGGG	936
QY	219	IleLysArgThrIleAlaSerLysValSerLeuGlyValaValTyrPhePheMetAsnGly	238
DB	937	ATPAAAGAAAGCTATCACGGCCCAACATTTCTATGGTGGCGCTTCTTATTAATGATGCA	996
QY	239	ThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluPheGlyTyr	258
DB	997	TCATATGCTCTGGCTTCTGGTATGGACCTCTCTGGTCTCTCCAGTGAA-----TAT	1050
QY	259	ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly	278
DB	1051	ACTATTGGACAGGTACTCACTGCTTCTTCTTCTGTTATTAATTTGGGGCTTTTAGTATGGA	1110
QY	279	AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe	298
DB	1111	CAGGCATCCCAAGCATTGAAGCATTTGCAACGCAAGAGGAGCAGCTTATGAATCTTC	1170
QY	299	GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu	318
DB	1171	AAGATAATTGACAATAAACCAAGCATTTGACAGCTATTTCGAAGAGTGCACATAACCAT	1230
QY	319	SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer	338
DB	1231	AATATTAAAGGAAATTTGGAATTTCAAAATGTTCACTTCAGTTACCCCTTCTCGAAAGAA	1290
QY	339	IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal	358
DB	1291	GTTAAGATCTTAAGGGTCTCAACCTGAAGTTTCAGAGTGGGCGAGCAGCTGGCGCTGGT	1350
QY	359	GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro	378
DB	1351	GGGAACAGTGGCTGCGGAAGAGCAGCAGCTGCAGCTGATGCAGAGGCTCTATGACCCC	1410
QY	379	AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr	398
DB	1411	ACAGTGGCATGCTGCTGTTATGATGGACAGGACATTAGGACCATTAATGTAAGCATCTT	1470
QY	399	ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn	418
DB	1471	CGGAAATTAATGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	1530
QY	419	AsnIleLysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGlu	438
DB	1531	AACATTCGCTATGGCCGCGAATAATGTCCATGGATGAGATTGAGAAAGCTGTTAAGGAA	1590
QY	439	AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu	458
DB	1591	GCCAAATGCTTATGATTTTATCATGAAACACTACCTTAATAAATTTGACACTCTGGTGGAG	1650
QY	459	LysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal	478
DB	1651	AGAGGGCCCGCTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	1710
QY	479	ArgAsnProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys	498
DB	1711	CGCAACCCCAAGATTCTTCTGCTGGATGAGCAACAGCTCAGCTCTGGACACTGAAGTGA	1770
QY	499	SerAlaValGlnAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAla	518
DB	1771	CGAGTGGTTCAGTGGCCCTGGATAGGCGAGAAAGGCGGAGCTACCATTTGTGATGCT	1830
QY	519	HisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeu	538
DB	1831	CATCGTTTGTACAGTTCGTATGCGATGTCATTCCTGCTGCTGCTGCTGCTGCTGCTGCT	1890
QY	539	AlaGluGlyGlyAlaHisAlaGluLeuMetAlaLeuArgGlyLeuTyrTyrSerLeuVal	558

D	b	1991	GTGAGAGAAAGAAATCATGATGAACCTCATGAAGAGAGAGGGCAGATTACTTCCAAACTTGTCT	1995
Q	y	559	MetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTySerThrGlu	578
D	b	1951	ACAATGCAG--ACAAGAGGAATGAAATTGAGTTAGAAAATGCCACTGGTGAATCCAAA	2007
Q	y	579	ArgLysThrAsnSerLeuProLeuHisSerValLysSerLleLysSerAspPheIle---	597
D	b	2008	AGTGAAGAGTGATGCTTGGAAATG--TCTCCAAAAGATTCAGGGTCCAGTTTAATAAAA	2064
Q	y	598	-----AspLysAlaGluGluSer	603
D	b	2065	AGAAGATCAACTCGCAGGAGTATACATGCACCACCAAGGCCAAGACCAAGACTTGGTACA	2124
Q	y	604	ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn	623
D	b	2125	AAAGAGGACTTGAATGAGATGTTCTTCTGAGGAGTTCGAAGCTGAAC	2184
Q	y	624	LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal	643
D	b	2185	TCAACTGAATGGGCTTATTTTGTGGTGGTATATTTTGTGCTATTATATAACGGAGGCCTG	2244
Q	y	644	HisProValPheSerIleIlePheAlaLysIleIleThrMetPhe--GlyAsnAsnAsp	662
D	b	2245	CAACGACGATTTTCAATAATATTTTCAAGGATATATAGGATCTTTACC CGGAGTAGGAT	2304
Q	y	663	LysThrThrLeuLysHisAspAlaGluIleTySerMetIlePheValIleLeuGlyVal	682
D	b	2305	CCTGAACAACAAACAGACAGATAGTACATGTTTTCTGTATGTTTCTAGTCTTGGAAAT	2364
Q	y	693	IleCysPheValSerTyPheMetGlnGlyLeuPheTyGlyArgAlaGlyGluIleLeu	702
D	b	2365	ATTTCTTTTATACATTTTCTCCAGGCGTTCCATATTTGGCAAAAGCTGGGAGATCCTC	2424
Q	y	703	ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyGlnAspIleAlaTrpPhe	722
D	b	2425	ACTAAGCGGCTTGATACATGGTTTTTCCAGTCCATGCTGAGACAGGATGTCCAGCTGTT	2484
Q	y	723	AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln	742
D	b	2485	GATGACCTTAAACACACCATTGGAGCATTGACACACGCGTTGCCAATGATCGCGCTCAA	2544
Q	y	743	IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly	762
D	b	2545	GTAAAGGGGCTATAGTTCACGCGTTGCTGTCAATCCACGAATATAGCAAACTCTTGGG	2604
Q	y	763	LeuSerValIleIleSerPheIleTyGlyTrpGluMetThrPheLeuIleLeuSerIle	782
D	b	2605	ACAGGCATTTATATATCTTAAATCTATGTTGGCATAAACACTTTTACTCTTACCAAT	2664
Q	y	783	AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn	802
D	b	2665	GTACCCATCATGTGCATAGCAGAGTGTGTAATGAAATGTTGCTGGACAAGCACTG	2724
Q	y	803	LysAspLysGlnLeuLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle	822
D	b	2725	AAAGATAAGAAAGAGCTAGAAGAGCTGGGAGATTGCTACAGAGCCATCGAAAACCTC	2784
Q	y	823	ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyGluGluMetLeu	842
D	b	2785	CGAACTGTGTTGTTCTTTGACTCGGGAGCAGAGTTTTGAATACATCATGTATGCACAGAGTTG	2844
Q	y	843	GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyAlaPhe	862
D	b	2845	CAAGTACCATACAGAAACTCTTTGAGAAGACACATCTCGGGGCTTCATTTTCTATC	2904
Q	y	863	SerHisAlaPheIleTyPheAlaTyAlaAlaGlyPheArgPheGlyAlaTyLeuIle	882
D	b	2905	ACCCAGGCAATGATGATTTTCTATGCTGGCTGTTTCCGGTTGGTGGCTACTTGGTG	2964
Q	y	893	GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyGly	902
D	b	2965	GCAATGAGTTCATGAACATTTCCAGGATGTTCTTTGGTATCTTCCAGTATTGCTTTGGT	3024

Qy	903	AlaValAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly	922
Db	3025	GCATGCGAGTGGGGCAGTCAGTTCATTTGCTCTGACTATGCCAAAGCCAAAGTATCA	3084
Qy	923	AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu	942
Db	3085	GCAGCCCACTCATATGATCATTTGAAATAAGCCCTCGATTGACAGCTACAGCGCCCTCAC	3144
Qy	943	GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr	962
Db	3145	GGCTCTAAGCCAAATACGTTGGAGAGAAATGTGACATTTAATAGGTCTGTTCAACTAT	3204
Qy	963	ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys	982
Db	3205	CCCACTCGACAGACATCCCGTGTCTCCAGGGCTGAGCCTCGAGGTGAAGAAGGCCAC	3264
Qy	983	ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln	1002
Db	3265	ACGTGTGCCCTCGTAGGTAGCAGTGGCTGTGGGAAGACACAGATTGTTTACGCTCTAGAG	3324
Qy	1003	ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu	1022
Db	3325	CGCTTCTATGACCCCTTGGCTGTCTCAGTGTCTAATTGATGGCAAGATAAACACCTG	3384
Qy	1023	AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn	1042
Db	3385	ANTGTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCTCAGGAGGCCATCCTCTTTGAC	3444
Qy	1043	CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu	1062
Db	3445	TGCAGCATTCGCGAGAACATTGCTATGGAGACAAACAGCGGGTCTATCATCATGAAGAG	3504
Qy	1063	IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys	1082
Db	3505	ATTATGCGACGCCACCAAGGAGGCCAACATACACCACTTCATCGAGACACTCCCTGAGAAA	3564
Qy	1083	TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu	1102
Db	3565	TACAACACAGATAGAGACAAAGAACCCAGCTCTCTGGTGGCCAGNAACAGCGCATT	3624
Qy	1103	AlaIleAlaArgAlaLeuLeuGlnLysProTyrIleLeuLeuLeuAspGluAlaThrSer	1122
Db	3625	GCCATAGCTCGCGCTCTTGTTAGACAGCCCTCATATTTTGTCTTTGGATGAAGCTACATCA	3684
Qy	1123	AlaLeuAspAsnAspSerGlnLysValValGlnHisAlaLeuAspLysAlaArgThrGly	1142
Db	3685	GCTCTGATACAGAAAGTGAAGAGTTGTCCAAAGAGCCCTGGACAAAGCCAGAGAAGGC	3744
Qy	1143	ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal	1162
Db	3745	CGCACCTGCAATTGTGATTCGCCACCGCTTGTCCACATCCAGATGCGAGATTTAATAGTG	3804
Qy	1163	ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg	1182
Db	3805	GTGTTTCAGAAATGGCAAGTCAAGGAGCATGCCACATCAACAGCTGCTGGCTCAGAAA	3864
Qy	1183	AspIleTyrPheLysLeuValAsnAlaGlnSer	1193
Db	3865	GGCATCTATTTTCCATGATCAGTGTCCAGGCT	3897

Qy	61	TyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLys	80
Db	487	TACATTCACGTTTCATTTTGGTCCTCGGCAGCTGGAAAGACAATACACAAAATTAGAAAA	546
Qy	81	GlnPhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGly	100
Db	547	CAGTTTTTTCATGCTAATAGGCACAGGAGATAGCTGGTTGATGTGCACGNTGTTGGG	606
Qy	101	GluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLys	119
Db	607	GAGCTTAACACCCGGCTTACAGATGATGTCCTCAAGATTAAATGAAGGAATTGGTCACAAA	666
Qy	120	IleAlaIleuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuVal	139
Db	667	ATTGGAATGTTCTTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATATAGGATTTACA	726
Qy	140	LysGlyTrpLysLeuThrLeuSerThrSerProLeuIleMetAlaSerAla	159
Db	727	CGTGGTTGGAGCTAACCTTGTGATTTTGGCATCAGTCCCTGTTCTTGGACTGTCACT	786
Qy	160	AlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLys	179
Db	787	GCAGTCTGGGCAAGATACTGTCTTCATTTACTGATAAAGAACTCTTAGCTTATGCAAAA	846
Qy	180	AlaGlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAla	199
Db	847	GCTGGAGCAGTAGCTCAAGAGGTCTTGGCGACAATTAGAACCTGTGATTCATTTGGAGGA	906
Qy	200	GlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIle	219
Db	907	CAAAAGAAAGAACTCGAAAGGTACAAACAAAATTTAGAAAGAGCTTAAAGAAATTCGGGATA	966
Qy	220	LysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThr	239
Db	967	AGAAGACTATTACACCCCAATATTTCTATAGTGTGCTGCTTCTCTTATCTATGCATCT	1026
Qy	240	TyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThr	259
Db	1027	TATGCTCTGGCCCTCTGGTATGGACCACTTGGCTCTCTCTCAAGGAA-----TATTCT	1086
Qy	260	IleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAla	279
Db	1081	ATTGGACAAAGTACTCACTGTATTTCTGTATTAAATTTGGGGCTTTTAGTGTTCGACAG	1144
Qy	280	AlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGln	299
Db	1141	GCATCTCCAGCATTTGCAAATTCGACAGGNGCAGCTTTTGCAAATCTTCAAG	1200
Qy	300	ValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSer	319
Db	1201	ATAATTGATAAATAAGCCAAGTATTGACAGCTATTTCGAAGAGTGGGCACAAACACGATAAT	1260
Qy	320	IleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIle	339
Db	1261	ATTAAAGGAAATTTGGAATTCGAATGTTCACCTACGTTACCCATCTCGAAAGAAAGTT	1320
Qy	340	LysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGly	359
Db	1321	AGAATCTTGAGGCGCTGAACCTGAAGGTGCAGAGTGGGCACACGGTGGCCCTGTGTGA	1380
Qy	360	LeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProSer	379
Db	1381	AACAGCGGCTGTGGGAAGAGCACACGGTCCAGCTGATGCAGAGGCTTTATGACCCACA	1440
Qy	380	AspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArg	399
Db	1441	GAGGCGCATGTGTGATGTGACAGGATATTAGGACCATTAACAGTAAAGTTTCTACGG	1500
Qy	400	AspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsn	419
Db	1501	GAATCATCTGGTGTGTGAGTCAGGAACCTGTATTGTTTGGCCACCACCATAGCTGAANAAC	1560
Qy	420	IleLysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAla	439

[illegible]

D	b	2635	GGAATAATTATATCTTAATCTATGTTGGCAACTGCACACTGTCTTACTCTTAGCAATTTGTA	2695
Q	y	784	ProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLys	803
D	b	2695	CCCATCTTCGAATCAGCGAGTGTTGAAATGAAAATGTTCTGTCTGGACAAGCACTGAAA	2754
Q	y	804	AspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArg	823
D	b	2755	CATAAGAAGAACACTAGAAGGTCGTGGGAAGATCGCTACTCTGAAGCAANTAGAAAACCTCCGA	2814
Q	y	824	ThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGln	843
D	b	2815	ACTGGTGTCTTGACTCAGGAGCAGAAGTTTGAAACATATGATGATCAGAGTTTGCAG	2874
Q	y	844	ThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSer	863
D	b	2875	GTACCATACAGAAACTCTTTTGAGGAAGCACACATCTTTTGGAAATCAGCTTTTTCTTCAG	2934
Q	y	864	HisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyValaTyrIleuIleGln	883
D	b	2935	CAGCAATATGATATTCTTCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGCCA	2994
Q	y	884	AlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAla	903
D	b	2995	CATAGTCTCATGAGCTTTGAGGATGTTCTGTAGTATTTTCAGCTGTGTCTTTGGTGCC	3054
Q	y	904	MetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAla	923
D	b	3055	ATGCCCTGGGCAAGTCAGTTTCATTTGCTCTCTGACTATGCCAAAGCCAAAGTATCAGCA	3114
Q	y	924	AlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGly	943
D	b	3115	GCCCACATCATCATGATCATTTGAAAAAACCCTTTGATTGACAGCTACAGCACAGAAGGC	3174
Q	y	944	LysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrPro	963
D	b	3175	CTAAGGCGCAACACATTTGGAAGGAATATGCACATTTAATGAAGTTGTATTCAACTATCCC	3234
Q	y	964	CysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThr	983
D	b	3235	ACCCGACTGGACATCCAGTGCTTCAGGGGCTGAGGCTCGGAAGTGAAGAAGSGCCAGAGC	3294
Q	y	984	ValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArg	1003
D	b	3295	CTGGCCCTGTGTGGCAGCAGTGCTGTGGGAAGACACGGTGCTCCAGCTCTCTGGAGCGG	3354
Q	y	1004	LeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsn	1023
D	b	3355	TTCATGACCCCTTTGGCGGGAAGTGCTCTTGAGCGCAAGAAATAAAGCACTGAAT	3412
Q	y	1024	ValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCys	1043
D	b	3415	GTTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCCAGGAGCCCCATCCTGTTGACTGC	3474
Q	y	1044	SerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIle	1063
D	b	3475	AGCATTAGTGAGAACATTGCTCATGGAGACAAACAGCGGGGTGTGTCTCAGGAGAGATC	3533
Q	y	1064	LysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyr	1083
D	b	3535	GTGAGGGCAGCCAAAGGAGGCCAATATACGCCCTTCATCGAGTCACTGCTTAATAAATAT	3594
Q	y	1084	AnThrGlnValGlyLeuLysGlyValaGlnLeuSerGlyGlyGlnLysGlnArgLeuAla	1103
D	b	3595	AGCACCAGATGAGAGCAAAAGNACTCAGCTCTCTGGTGGCCAGAAACACGCATTGCC	3653
Q	y	1104	IleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAla	1123
D	b	3655	ATAGCTGTGCCCTTGTATAGACGCCTCATATTTTCTTTTGGATGAAGCCACATCAGCT	3712
Q	y	1124	LeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArg	1143
D	b	3715	CTGGATACAGAAAGTGAAGAGTTGTCCAAAGAGCCCTCGCAAGAGCCAGAGAAGGCCGT	3774

Qy	241	GlyLeuAlaPheTyrGlyThrSerLeuIleLeuAenGlyGluProGlyTyrThrIle	260
Db	931	GCTCTGGCCCTTCGGTATGGACACCTTGGTCTCTCAGGGAA-----TATTCTATT	984
Qy	261	GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla	280
Db	985	GGACAAGTACTCACTGTATCTCTTTCTGTATTAATGTGGGCTTTTAGTGTGGACAGCA	1044
Qy	281	ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGluVal	300
Db	1045	TTCTCCAAAGCATTGAAGCATTTGCAATGTCAGAGGAGCAGCTTATGAATCTTCAGATA	1104
Qy	301	IleAspLysLysProSerIleAspAenPheSerThrAlaGlyTyrLysProGluSerIle	320
Db	1105	ATTGATAATAAGCCAGTATTGCACGTATTCGAGAGTGGGCACAAACACAGATAAATT	1164
Qy	321	GluGlyThrValGluPheLysAenValSerPheAenTyrProSerArgProSerIleLys	340
Db	1165	AAGGMAATTTGGAAATTCAGMAATGTTCACCTTCAGTTACCCATCTCGAAAAGAAATTAA	1224
Qy	341	IleLeuLysGlyLeuAenLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu	360
Db	1225	ATCTTGAAGGCCCTGSAACCTGAAGTGCAGAGTGGCAGACGGTGGCCCTGGTTGGAAC	1284
Qy	361	AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp	380
Db	1285	AGTGGCTGTGGGAAGACACACAGTCCACGTATGTCAGAGGCTCTATGACCCACAGAG	1344
Qy	381	GlyPheIleMetValAspGluAenAspIleArgAlaLeuAenValArgHisTyrArgAsp	400
Db	1345	GGGATGTCTAGTCTTGATGCACAGGATATTAGGACCATAAATTAAGGTTTCTACGGAA	1404
Qy	401	HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle	420
Db	1405	ATCATTTGCTGTGTAGTCAGGAACCTGTATTGTTGGCCACACGATGCTGAAACCAT	1464
Qy	421	LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn	440
Db	1465	CGTATGGCCGTGAAATGTCACCATGGATGAGATTGAGAAAGCTGTACAGAAGCCAT	1524
Qy	441	AlaTyrAspPheIleMetGluPheProAenLysPheAenThrLeuValGlyLysGly	460
Db	1525	GCCTATGACCTTATCATGAACTGCCTCATAAATTTGACACCCCTGTTGGAGAGAGGG	1584
Qy	461	AlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn	480
Db	1585	GCCAGTTGAGTGTGGCAGACAGAGGATCGCCATTTCACGTCGCCCTCGTTGCGAAC	1644
Qy	481	ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla	500
Db	1645	CCCAAGATCTCTCTGTGGATGAGGCCAGCTCAGCCTTGGACAGAAGGCGAAGCATGT	1704
Qy	501	ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg	520
Db	1705	GTTTCAGGTGGCTCTGGATAGGCCAGAAAGGTTCGACCCACCATTTGATAGCTCATCGT	1764
Qy	521	LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu	540
Db	1765	TTGTCTACAGTTCTGTAATGTGACGTCATCGCTGTTTTCGATGATGGAGTCAATTGGAG	1824
Qy	541	LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer	560
Db	1825	AAAGGAATCATGATGAACATCAAGAAGAAGGCAATTTACTTCAAACATGTGCACAAATG	1884
Qy	561	Gln-----AspIleLysLysAlaAspGluGlnMetGluSerMetThr	574
Db	1885	CAGACAGCAAGAAATGAGTTGAATGTAGAAATGCAAGCTGATGAATCCAAAAGTCAAATT	1944
Qy	575	TyrSerThrGluArgLysThrAenSerLeuProLeuHisSerVal-----	589
Db	1945	GATCCCTTGGAAATGTCTTCAAATGATTCAAAGTCCAGTCTAATAAGAAAAAGATCAACT	2004

Qy	590	--LysSerIleLys---SerAspPheIleAspLysAlaGluSerThrGlnSerLys	607
Db	2005	CGTAGGAGTGTCCGTGCATCAAAAGCCCAAGCAAAAGCTTAGTACCAAGAGCGCTCTG	2064
Qy	608	GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTirp	627
Db	2065	GATGAAAGTATACCTCCAGTTTCCTTTGGAGGATATGAAGCTAAATTTAACTCAATATGG	2124
Qy	628	ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe	647
Db	2125	CCATATTTGTTGTTGGTGTATTTTGTGCCATTAAATGAGGCGCTCAACACGACATTT	2184
Qy	648	SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu	666
Db	2185	GCAATAATATTTCAAAGATTATAGGGGTTTTTACAAGAAATGATGATCCTGCAACAAAA	2244
Qy	667	LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal	686
Db	2245	CGACAGAATAGTAACTGTTTTCACATATTGTTTCAGGCCCTGGAAATTAATTTCTTTTATT	2304
Qy	687	SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu	706
Db	2305	ACATTTTTCCTTCAGGTTTTCACATTTGCGAAAGCTGAGAGATCTCTCACCAGGGCTC	2364
Qy	707	ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTirpPheAspGluLysGlu	726
Db	2365	CGATACATGTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTGTATGACCTAAA	2424
Qy	727	AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla	746
Db	2425	AACACCACTCGAGCATTTGACTACAGGCTCGCAATGATGCTGCTCAAGTTTAAAGGGCT	2484
Qy	747	ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle	766
Db	2485	ATAGTTTCAGGCTTGCTGTAAATCCAGATATAGCAATCTTGGACAGGAATAATT	2544
Qy	767	IleSerPheIleTyrGlyTirpGluMetThrPheLeuIleLeuSerIleAlaProValLeu	786
Db	2545	ATATCCTTCATCTATGTTGGCAACTTAACACTGTTTACTCTTAGCAATTTGATCCCATCAT	2604
Qy	787	AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln	806
Db	2605	GCAATACAGCAGGTTGTTGAAATGAAAAATGTTGTTCTGGCAAGCACTCAAAAGATAAGAA	2664
Qy	807	GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal	826
Db	2665	GAACTAGAGGTGCTGGAGAGATCGCTACTGAAGCAATAGAAAACTTCGGAACCGTGTGT	2724
Qy	827	SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis	846
Db	2725	TCCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTCCAGGTACCATAC	2784
Qy	847	ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe	866
Db	2785	AGAAACTCTTTTGAGGAAGCACACATCTTTTGGAAATTTACATTTTCTCCACCAGGCAATG	2844
Qy	867	IleTyrPheAlaValAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg	886
Db	2845	ATGTATTTTTCCTATGCTGAGATGTTTCCGGTTTGGAGCCCTACTTGGTGGCACAATAACTC	2904
Qy	887	MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle	906
Db	2905	ATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTGTGTTCTTGGTGCCATGCCGTG	2964
Qy	907	GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaHisLeu	926
Db	2965	GGGCAAGTCAGTTTCATTTGCTCTGACTATGCCAAGCCAAAATATCATGAGCCGCACATC	3024
Qy	927	PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro	946
Db	3025	ATCATGATCATTTGAAAAAACCCTTTGATTGACAGCTACAGCACCGGAGGCCCTAATGCCG	3084
Qy	947	AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro	966


```
Db 3085 AACACATTGGAAGAAATGTCACATTGGTGAAGTTGTTATTCACATCCACCCGACCG 3144
Qy 967 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLeuValAlaPhe 986
Db 3145 GACATCCCATGCTGTTCAGGACCTGAGCCTGGAGTGAAGAAGGCCAGACGCTGGCTCTG 3204
Qy 987 ValGlySerSerGlyCysGlyLysSerThrValGlnLeuLeuGlnArgLeuTyrAsp 1006
Db 3205 GTGGCAGCAGTGGCTGTGGAGAGACACAGTGTCCAGTCTCTGGAGCGTCTACGAC 3264
Qy 1007 ProValGlnGlnValLeuPheAspGlyValAlaAspAlaLysGluLeuAsnValGlnTrp 1026
Db 3265 CCCTTGGCAGGAAAGTGTCTGTGTCGCAAGAAATAAAGCGACTGAATGTTCACTGG 3324
Qy 1027 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 1046
Db 3325 CTCGAGACACCTGGGCATCGTGTCCAGAGCCCATCTCTTTCAGTGCAGCATGCT 3384
Qy 1047 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 1066
Db 3385 GAGAACATTCCTATGGAGACAACAGCCGGTGTGTACAGGAAGATCGTGAGGCA 3444
Qy 1067 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 1086
Db 3445 GCAAAGAGGCCAACATACATGCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAA 3504
Qy 1087 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 1106
Db 3505 GTAGGAGACAAAGAACTCAGCTCTCTGGTGGCCAGAAACAGCATTCGCTAGCTCGT 3564
Qy 1107 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 1126
Db 3565 GCCTTGTAGACAGCTCATATTTTGTGTTGATGAAGCCAGCTCAGCTCTGGATACA 3624
Qy 1127 AspSerGluLysValValGlnHisAlaLeuAspLysAlaAlaArgThrGlyArgThrCysLeu 1146
Db 3625 GAAAGTGAAAGGTTGTCCAAGAAAGCCCTGGACAAGCCAGAGAGCGCCACCTGCATT 3684
Qy 1147 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1166
Db 3685 GTGATTGCTCAGCGCTGTCCACCATCCAAATGCAGACTTAATAGTGGTGTTCAGAAAT 3744
Qy 1167 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1186
Db 3745 GCAGAGTCAAGAGCATGTGCAGCATCAGCTGTCTGGCACAGAAAGCACTATTTT 3804

Qy 1187 LysLeuValAsnAlaGlnSer 1193
ID ABA94366
XX
AC ABA94366;
XX
DT 26-MAR-2002 (first entry)
XX
DE Human BCRP DNA related seq id No. 3.
XX
KW Stem cell; ATP transport protein; ATP-binding cassette; antiparkinsonian;
KW hepatocytic; neurodegenerative; cytostatic; antianemic; muscular; BCRP;
KW cardiac; gene therapy; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..3843
FT /*tag= a
XX
XX WO200192877-A2.
```

```
PD 06-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US17459.
XX
PR 31-MAY-2000; 2000US-0584586.
PR 29-MAY-2001; 2001US-0866866.
XX
PA (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX
PI Sorrentino B, Schuetz J;
XX
XX WPI; 2002-114368/15.
DR P-PSDB; ABB07267.
XX
PT Identifying a stem cell, for treating e.g., muscular dystrophy,
PT myocardial infarction, Parkinson's disease, or neurodegenerative
PT disorders, comprises detecting the expression of an ATP transport
PT protein (BCRP) by a cell -
XX
PS Disclosure; Page 59-60; 87pp; English.
XX
CC The invention provides a method of identifying and/or isolating a stem
CC cell that involves detecting the expression of an ATP transport protein
CC containing a conserved ATP-binding cassette (BCRP) by a cell in a sample
CC comprising stem cells. The isolated stem cells may be used in the
CC treatment of diseases such as muscular dystrophy, degenerative liver
CC disorder, myocardial infarction, Parkinson's disease, degenerative liver
CC disorders of the brain, and for tissue regeneration or replacement.
CC Haematopoietic cells can be used in bone marrow transplants (e.g., for
CC treatment of leukemia) and for ex vivo gene therapy for treating blood
CC diseases such as sickle cell anemia and thalassemia. The stem cells can
CC also be used as cell targets in gene therapy protocols. The present
CC sequence represents a sequence related to the BCRP for which no relevant
CC information has been provided in the specification.
XX
SQ Sequence 3860 BP; 1135 A; 746 C; 956 G; 1023 T; 0 other;

Alignment Scores:
Pred. No.: 0 Length: 3860
Score: 2545.00 Matches: 681
Percent Similarity: 75.55% Conservatives: 246
Best Local Similarity: 55.50% Mismatches: 264
Query Match: 58.97% Indels: 36
DB: 24 Gaps: 9

US-09-873-409-6 (1-1195) x ABA94366 (1-3860)
Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 151 ATGGTGGTGGAACTTTGGCTGCCATCATCCATCGGGCTGGACTTCCTCTCATGATGCTG 210
Qy 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
Db 211 GTGTTGGAGAAATGACAGATATCTTTGCAATGTCAGGAAATTTAGAGATCTGATGTCA 270
Qy 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
Db 271 AACATCACTAATAGAGATGATATCATGATACAGGGTTCTTCATGATCTGGAGGAGAC 330
Qy 44 -----ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db 331 ATGACCAGATATGCTTATTATACAGTGGAAATTTGGTGGGTGCTGGTGTCTGCTTAC 390
Qy 62 IleGlnIleSerLeuTrpIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db 391 ATTCAGTTTCAATTTTGGTGCCTGGCAGCTGGAAGACAAATACAAAAATTACAAACAG 450
Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101
Db 451 TTTTTCATGCTAATATGACAGGAGATAGGCTGGTTGATGTGTCACCATGCTTGGGAG 510
Qy 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
```


511	Db	CTTAAACACCCGACTTACAGATGATGCTCTAGATTAATGAAGTATTATGGTGACAAAATT	570
121	Qy	AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys	140
571	Db	GGAAATGTTCTTTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGGATTTACACGT	630
141	Qy	GlyTyrPheLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla	160
631	Db	GGTTTGAAGCTAACCCCTGTGTGATTTTGGCCATCAGTCTCTGTTTCTTGGAGCTGCTGAGCTGCT	690
161	Qy	AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla	180
691	Db	GTCTGGGCAAGATACATCTCTCACTTACTGATAAGAACTCTTAGCGTATGCAAAAGCT	750
181	Qy	GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln	200
751	Db	GGAGCAGTAGCTGAAGAGGCTCTGGCAGCAATTAGAACTGTGATTCATTTGGAGGACAA	810
201	Qy	GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyLysLys	220
811	Db	AAGAAAGAACTTTGAAGAGGTACAAACAAAATTTAGAAAGAGCTTAAAGAAATTTGGGATAAAG	870
221	Qy	ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyTyrTyr	240
871	Db	AAAGCTATTACAGCCAAATATTTCTATAGGTGCTGCTTCTCTGCTGATCTATGATCATCTTAT	930
241	Qy	GlyLeuAlaPheTyrPyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle	260
931	Db	GCTCTGGCCTTCTGGTATGGACCACTTGGTCTCTCTCAGGGGAA-----TATCTTATT	984
261	Qy	GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla	280
985	Db	GGACAAGTACTCACTGTATCTTTCTGTATTAAATTGGGGCTTTTAGTGTGGCAGGCA	1044
281	Qy	ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal	300
1045	Db	TCCTCAAGCAATTGAAGCATTTTGCAATGCAAGAGAGAGAGCTTTATGAAATCTTCACAGATA	1104
301	Qy	IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle	320
1105	Db	ATTGATAATAAGCCAAGTATTGCACAGCTATTGCAAGAGTGGGCACAAACACGATAAATATT	1164
321	Qy	GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys	340
1165	Db	AAGGAAATTTTGGAAATTCAGAAATGTTCATCTTCACTTACCCTCTCGAAAAAGAGTTAAG	1224
341	Qy	IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu	360
1225	Db	ATCTTGAAGGCCCTGAACCTGAAAGTGGCAGAGTGGGCACGCGTGGCCCTGGTTGGAAAC	1284
361	Qy	AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp	380
1285	Db	AGTGGCTGTGGGAAGAGCAACAAGTCCAGCTGTATGCAAGAGGCTCTATGACCCCAACAGAG	1344
381	Qy	GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp	400
1345	Db	GGCATGTCTAGTCTTCATGCACAGGATATTAGCAACATAAATAGAGTTTCTACGGGAA	1404
401	Qy	HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle	420
1405	Db	ATCATTTGGTGTGTAGTCAGGAACCTGTATTGTTTGGCCACCAAGATAGCTGAAATAACATT	1464
421	Qy	LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn	440
1465	Db	CGCTATGGCCGTGAAATGTCCACCATGGATGAGATTGAGAAGAGTGTCAAGGAAGCCAAAT	1524
441	Qy	AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGly	460
1525	Db	GCCTATGACTTTATCATGAAACTGGCTCATAAATTTTGACACCTTGGTGGAGAGAGGG	1584
461	Qy	AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn	480
1585	Db	GCCCAAGTGTGGTGGGAGAGAGAGAGATGCGCATTTGCACTGCGCTGTTGCGAAC	1644

Qy	481	Prol	ysileuileuileuaspGlu	alathrSerAlaLeuAaspSerGluSer	ysSerAla	500
Db	1645	CCCAAGATCCTCGTCTGATGAGCCACGTCAGCGCTTGGACACAGAANAAGCGAAGCAAGT	1704			
Qy	501	ValGlnAlaAlaLeuGlu	lysAlaSer	lysGlyArgThrThrIleValValAlaHisArg	520	
Db	1705	GTTTCAGGTGGCTCTCGATTAAGGCCAGANAAGGTCGACACCACTTGTGTATGATGCTCATCGT	1764			
Qy	521	LeuSerThrIleArgSerAlaAsePhe	leuIleValThrLeuLysAaspGlyMetLeuAlaGlu	540		
Db	1765	TTGTCTACAGTTCTGTAATGCTGACGTCATCGCTGGTTCGATGATGGAGTCAATTGTGTGGAG	1824			
Qy	541	LysGlyAlaHisAlaGluLeuMetAla	lysArgGlyLeuTyrTyrSerLeuValMetSer	560		
Db	1825	AAAGGAAATCATGATCAACTCATGAAGAAGAACGATTTACTTCAAACTTTGTCCACAATG	1884			
Qy	561	Gln-----Asp	leuLysAlaAaspGluGlnMetGluSerMetThr	574		
Db	1885	CAGACAGCAGGAATCAAGTTGAATTAGAAAATGACGCTGATGAATCCAAAAGTGAAATT	1944			
Qy	575	TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal-----	589			
Db	1945	GATGCTTGGAAATGCTTCAAAATGATTCAAGATCCAGTCTAATAAGAAAAGATCAACT	2004			
Qy	590	---LysSerIleLys---Ser	AspPheIleAePlysAlaGluSerThrGlnSerLys	607		
Db	2005	CGTAGGAGTGTCCGTGGATCACAAAGCCCAAGACAAAGCTTAGTACCCAAGAGCGCTCTG	2064			
Qy	608	GluIleSerLeuProGluValSerLeu	leuLysIleLeuLysLeuAsnLysProGluTyr	627		
Db	2065	GATGAAGATATACCTCCAGTGTTCCTTTTGGAGGATTAATGAAGCTTAAATTTAACTCGAATGG	2124			
Qy	628	ProPheValValLeuGlyThrLeuAlaSerValLeu	AsnGlyThrValHisProValPhe	647		
Db	2125	CCTTATTTGTGTGGTGATTTTGTGCCATTAATAATGAGGCGCTGCACACAGCATTT	2184			
Qy	648	SerIleIlePheAlaLysIleIleThrMetPhe	GlyAsn---AsnAspLysThrThrLeu	666		
Db	2185	GCAATAATATTTTCAAGATTTATAGGGTTTTTTTCAAGAAATTTGATGATCCTGAAACAAA	2244			
Qy	667	LysHisAaspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal	686			
Db	2245	CGACAGAATAGTAACCTGTGTTTCTACTATTGTTCTAGGCCCTTGGAAATTTATTTCTTTTATT	2304			
Qy	687	SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu	706			
Db	2305	ACATTTTCTTCAGAGTTTCACATTTGGCAAGCTGGAGAGATCCTCACCACAGCGGCTC	2364			
Qy	707	ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGluLysGlu	726			
Db	2365	CGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTGATGACCCATAA	2424			
Qy	727	AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAaspIleAlaGlnIleGlnGlyAla	746			
Db	2425	AACACCACTGGACATTGACTACCAGGCTCGCCAATGATGTGCTCAAGTTTAAAGGGGCT	2484			
Qy	747	ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle	766			
Db	2485	ATAGGTTCCAGGGTCTGCTGTAATATCCCAAGATATAGCAAAATCTTGGGACAGGAATAATT	2544			
Qy	767	IleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeu	786			
Db	2545	ATATCCTTCACTATGTTGGCACTTAACACTGTTTACTTCTTAGCAATTTGTACCCATCATTT	2604			
Qy	787	AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln	806			
Db	2605	GCAATGACAGGAGTTGTTGAAATGAAAATGTTGTCTGCAAGCACTGCARAAGATAAGAA	2664			
Qy	807	GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal	826			
Db	2665	GAACTAGAAGGTCTGGGAAGATCGCTACTGAAGCAATAGAAAACCTTCGGAACCGCTGTT	2724			


```
QY 827 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGlnMetLeuGlnThrGlnHis 846
DB 2725 TCCTTGACTCAGGACGAGAAAGTTTGAACATATGATGCTCAGAGTTTGCAGGTACCATAC 2784
QY 847 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 866
DB 2785 AGAAACTCTTTGAGGAAGACACATCTTTGGGAATATACATTTTCCCTTCACCCAGGCAATG 2844
QY 867 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 886
DB 2845 ATGTAATTTTCCATGCTGATGTTCCGGTTTGGAGCCTACTTGTGGCACAATAAATC 2904
QY 887 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 906
DB 2905 ATGAGCTTTCAGGATGTTCTGTAGTATTTCAGTCTGTCTTGTGTCATGGCCGCTG 2964
QY 907 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 926
DB 2965 GGGCAAGTCAGTTCAATTTGCTCTGACTATGCCAAAGCCAAATATATCAGCAGCCACATC 3024
QY 927 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 946
DB 3025 ATCATGATCATTTGAAAACCCCTTTGATTTGACAGCTACAGCAGCGAGGCTTAATGCCG 3084
QY 947 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 966
DB 3085 AACACATTGGAAGGAATGTACATTTTGGTGAAGTTGATTTCAACTATCCACCCGACCG 3144
QY 967 AspValPheIleLeuArgGlyLeuSerIleLeuArgGlyLysThrValAlaPhe 986
DB 3145 GACATCCAGTGTCTTCCAGGACTGAGCTCGAGGTGAAGAAGCCAGCAGCTGGCTCTG 3204
QY 987 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 1006
DB 3205 GTGGCAGCAGTGGCTGTGGAGAGACACAGTGTCCAGCTCTCTGGAGCGGTCTTACGAC 3264
QY 1007 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 1026
DB 3265 CCCTTGGCAGGAAAGTGTCTGTATGTTGTCGAAATAAAGCGACTGAATGTTCACTGG 3324
QY 1027 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 1046
DB 3325 CTCGAGCACACCTGGGCATCGTGTCCAGGAGCCCATCTCTGTTGACTGCACGATGCT 3384
QY 1047 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 1066
DB 3385 GAGAACATTGCTATGGAGACACAGCCGGTGTGTACAGGAGAGATCGTGGGGCA 3444
QY 1067 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 1086
DB 3445 GCAAGGAGGCCAACATACATGCTTCTATCGAGTCACTGCTTAATAATATAGCACTAA 3504
QY 1087 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAArg 1106
DB 3505 GTAGGACAAAGAACTCAGCTCTCTGTTGGCCAGAAACAGCCATTCGATGCTCGT 3564
QY 1107 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 1126
DB 3565 GCCCTGTTAGACAGCTCATATTTGCTTTGGATGAGCCAGCTGCTGAGCTGGAATACA 3624
QY 1127 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1146
DB 3625 GAAAGTGAAAGGTTGTCCAAAGAGCCCTGGACAAAGCCAGAGAGCCGACCTGCATT 3684
QY 1147 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuValValLeuHisAsn 1166
DB 3685 GTGATTGCTCACCCTGCTCCACCATCCAGATGCGAGCTTAATAGTGGTGTTCAGAAT 3744
QY 1167 GlyLysIleLysGlnGlnGlyThrHisGlnGlnLeuLeuArgAsnArgAspIleTyrPhe 1186
DB 3745 GGCAGAGTCAAGAGCATGTGCAGCATCAGCAGCTGTGTCACAGAAAGGCATCTATTTT 3804
QY 1187 LysLeuValAsnAlaGlnSer 1193
```

```
DB 3805 TCAATGTCAGTGTCCAGGCT 3825
RESULT 13
AAH57442
ID AAH57442 standard; cDNA; 4349 BP.
XX AAH57442;
XX 10-SEP-2001 (first entry)
XX Human intestine cell specific cDNA sequence SEQ ID NO:282.
XX Human; tissue specific; diagnosis; brain; heart; skeletal muscle;
KW lung; liver; uterus; ovary; stomach; intestine; kidney; pancreas; ss;
KW metabolic disease; developmental disease; cytostatic; immunomodulatory;
KW neuroprotective; gene therapy; cancer; immunopathology; neuropathology.
XX Homo sapiens.
XX WO200132927-A2.
XX 10-MAY-2001.
XX 02-NOV-2000; 2000WO-US30396.
XX 04-NOV-1999; 99US-0163508.
XX (INCY-) INCYTE GENOMICS INC.
XX Sornasse T, Seilhamer JJ, Watson GA;
XX WPI; 2001-291057/30.
XX New cell and tissue specific polynucleotides useful for diagnosis, is
PT prognosis or monitoring of treatments for disorders where the gene is
PT associated with a cancer, immunopathology or neuropathology -
XX Claim 1; Page 207-208; 327pp; English.
XX AAH57161 to AAH57576 represent cell and tissue specific polynucleotide
CC sequences (I). (I) can have cytostatic, immunomodulatory and
CC neuroprotective activities, and can be used in gene therapy. (I) and
CC proteins (II) encoded by then are used in high throughput screening
CC assays to select DNA molecules, RNA molecules, peptide nucleic acids,
CC mimetics, peptides, proteins, agonists, antagonists, antibodies or
CC their fragments, immunoglobulins, inhibitors, drug compounds and
CC pharmaceutical agents. Expression of (I) in a sample indicates the
CC differentiation of embryonic stem cells into a tissue selected from
CC brain, heart, kidney, liver, lung, skeletal muscle or pancreatic
CC tissues. (I) and (II) are used to produce an expression profile that
CC defines a metabolic or developmental process, treatment, condition,
CC disease or disorder. The gene profile can be used for diagnosis,
CC prognosis or monitoring of treatments and for investigating a
CC predisposition to a disorder where the gene is associated with a
CC cancer, immunopathology or neuropathology.
XX Sequence 4349 BP; 1304 A; 816 C; 1048 G; 1181 T; 0 other;
SQ
```

```
Alignment Scores:
Pred. No.: 0 Length: 4349
Score: 3545.00 Matches: 681
Percent Similarity: 75.55% Conservative: 246
Best Local Similarity: 55.50% Mismatches: 264
Query Match: 58.97% Indels: 36
DB: 22 Gaps: 9
US-09-873-409-6 (1-1195) x AAH57442 (1-4349)
QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
DB 276 ATGGTGGTGGGAACATTTGGCTGCCATCATCGGGCTGGAGCTTCTCTCATGATGCTG 335
```


Qy	21	ValLeuGlyGluMetSerAsp	-----AsnLeuIleSer	31
Db	336	GTGTTTGGAGAAATGACAGATATCTTTTGCAAATGCAGGAAATTTAGAAAGATCTGATGTCA	395	
Qy	32	GlyCysLeuValGlnThr	-----AsnThrTyrSerPhePhe	43
Db	396	AACATCACTAATAGAGTGTATCAATGATACAGGGTCTCTCATGAATCTCGAGGAGAC	455	
Qy	44	-----ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr	61	
Db	456	ATGACACAGGTATGCCATTATTACAGTGGAAATGGTCTGGGGTCTGGTGGCTGTAC	515	
Qy	62	IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln	81	
Db	516	ATTCCAGGTTCATTTTGGTGCCTCGGACGCTGGAGACAAATACACAAAATTTAGAAAAACAG	575	
Qy	82	PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu	101	
Db	576	TTTTTTTCATGCTATATATGCGACGAGGATAGCTGTGTGATGTCACAGTGTGGGGAG	635	
Qy	102	LeuAsnThrArgMetThr	---AspIleAspLysIleSerAspGlyIleGlyAspLysIle	120
Db	636	CTTAACACCCGACTTACAGATGATGCTCTAAGATTAAATGAAGTTATTGGTGACAAAATT	695	
Qy	121	AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValIys	140	
Db	696	GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATATAGGATTTACACGT	755	
Qy	141	GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla	160	
Db	756	GGTTGGAAAGCTAACCCCTTGATTTTGGCCATCAGTCCGTGTCTTGGACTGTGAGTGTCT	815	
Qy	161	AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla	180	
Db	816	GTCTGGCCAAAGATACTATCTTCAATTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT	875	
Qy	181	GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln	200	
Db	876	GGAGCAGTAGCTGAAGAGGTCTTGGCGCAATTAGAACTGTGATTGCATTTGGAGGACAA	935	
Qy	201	GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleIys	220	
Db	936	AAGAAAGAACTTGAAGAGGTACAAACAAAATTTAGAAAGAGCTTAAAGAAATTTGGGATAAAG	995	
Qy	221	ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyTyrTyr	240	
Db	996	AAAGCTATTACAGCAATATTCTATAGGTGCTGCTTCTCTGCTCATCTATGCACTCTAT	1055	
Qy	241	GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle	260	
Db	1056	GCTCTGGCCCTCTCGTATGGACACCTTTGGTCTCTCAGGGGAA-----TATTCTATT	1109	
Qy	261	GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla	280	
Db	1110	GGACAAGTACTCACTGATTCTTTTCTGTATTAAATTTGGGGCTTTTAGTGTGGCAGGCA	1169	
Qy	281	ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal	300	
Db	1170	TCTCCAGCATTTGCAATTTGCAATGCAAGGAGCGACTTATGAATCTTCAAGATA	1229	
Qy	301	IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle	320	
Db	1230	ATTGATTAATAGCCAAAGTATTGACAGCTATTCCAGAGAGTGGGCACAAACCATGATAATTT	1289	
Qy	321	GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleIys	340	
Db	1290	AAGGGAAATTTGGAATTCGAAATGTTCTCACTTCACTTACCCTACCCATCTCGAAAAAGAGTTAAG	1349	
Qy	341	IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu	360	
Db	1350	ATCTTGAAGGCCCTGAACCTGAAGGTGCAGAGTGGCGACGGTGGCCCTGGTGGAAAC	1409	
Qy	361	AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnAtgLeuTyrAspProAspAsp	380	

Db	1410	AGTGGCTGTGGGAAGACCAACAGTCCAGCTGATGTCAGAGGCTCTATGACCCACAGAG	1469
Qy	381	GlyPheIleMetValAspGluAenAspIleAArgAlaLeuAenValAtrGHiSerAsn	400
Db	1470	GGATGGTCAAGTGTGATGTCAGACAGATATTAGACCAATAAATGTAAGCTTTCTACGGGAA	1529
Qy	401	HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle	420
Db	1530	ATCATTTGGTGTGATGAGTCAGGAACCTGTATTGTTGTCACACGATAGCTGAACACATT	1589
Qy	421	LysTyrcIyArgAspAspValThrAspGluGluMetGluAArgAlaAArgGluAlaAsn	440
Db	1590	CGCTATGGCCGTGAAAATGTCACATGGATGAGATTGAGAAGCTGTCAAGGAAGCCCAAT	1649
Qy	441	AlaTyArgPheIleMetGluPheProAenLysPheAenThrLeuValGlyGluLysGly	460
Db	1650	GCCTATGACTTATCATGAACCTGCTCATAAATTTGACACCTGGTTGGAGAGAGAGGG	1709
Qy	461	AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaAArgAlaLeuValArgAsn	480
Db	1710	GCCCAAGTTGAGTGTGGGCAGGAAGCAGAGAGATCGCCAATTGCAGCTGCCCTGGTTGCGAAC	1769
Qy	481	ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAAspSerGluSerLysSerAla	500
Db	1770	CCCAAGATCCTCTGCTGTGATGAGGCCACGTCAGCTTTGGACACAGAAAGCGAAGCAGTG	1829
Qy	501	ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg	520
Db	1830	GTTCAGTGGCTCTGATGAAGCCAGAAAGCTCGACACACATTTGTGATAGCTCATCGT	1889
Qy	521	LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu	540
Db	1890	TTGTCTACAGTTCGTAATGCTGACGCTCATCGCTGTTGATGATGAGATCATTTGTGGAG	1949
Qy	541	LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyTySerLeuValMetSer	560
Db	1950	AAAGGAATCATGATCAACTCATGAAGAGAAAGACATTTACTTCAAACCTTGTCCACAATG	2009
Qy	561	Gln-----AspIleLysAlaAspGluGlnMetGluSerMetThr	574
Db	2010	CAGACAGAGGAANTGAAGTTGAATTAGAAAAATGACGCTGATGAATCCAAAAGTGAAATT	2069
Qy	575	TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal-----	589
Db	2070	GATGCTTGGAAATGCTTCAAAATGATTCAAGATCCAGCTCTAATAAGAAAAAGATCAACT	2129
Qy	590	---LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys	607
Db	2130	CGTAGAGAGTTCGCTGGATCACAAGCCCAAGACAGAAAGCTTAGTACCAGAGAGGCTGTG	2189
Qy	608	GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAenLysProGluTrp	627
Db	2190	GATGAAGATATACCTCCAGTTTCCTTTGGAGGATATGAAGCTAAATTTAACTGAATGG	2249
Qy	628	ProPheValValLeuGlyThrLeuAlaSerValLeuAenGlyThrValHisProValPhe	647
Db	2250	CCITATTTTGTGTTGTTATTTTGTGCCATTATAAATGGAGGCTCCACACGACATTT	2309
Qy	648	SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu	666
Db	2310	GCAATATATTTTCAAAGATATATAGGGGTTTTTACAGAAATTTGATGATCCTGGAACAAAA	2369
Qy	667	LysHisAspAlaGluIleTySerSerMetIlePheValIleLeuGlyValIleCysPheVal	686
Db	2370	CGACAGAAATGTAACTTGTTTTCATCTATTGTTTCTAGGCCCTTGGAATTAATTTCTTTATT	2429
Qy	687	SerTyxPheMetGlnGlyLeuPheTyArgAlaGlyGluIleLeuThrMetArgLeu	706
Db	2430	ACATTTTCTTCAGGGTTTCATTTGGCAAAGCTGGAGAGATCCTCACCACAGCGCTC	2489
Qy	707	ArgHisLeuAlaPheLysAlaMetLeuTyGlnAspIleAlaIatrpPheAspGluLysGlu	726

Db 2490 CGATACATGTTTCCGATCCATGCTCAGACAGGATGTGATGGTTGGTTGATGACCCCTAAA 2549
QY 727 AasnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyVala 746
Db 2550 AACACCACTGGAGCATTTGATACACAGCTCGCCAATGATGCTCAAGTTAAAGGGCT 2609
QY 747 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAenMetGlyLeuSerValile 766
Db 2610 ATAGGTTCCAGGCTGTGTAAATTACCCAGATATAGCAAAATCTTGGGACAGGAATAATT 2669
QY 767 IleSerPheIleThrGlyTyrGluMetThrPheIleLeuSerIleAlaProValLeu 786
Db 2670 ATATCTTCATCTATGTTGGCAATACCACTGTATCTTACCAATGTACCCATCAT 2729
QY 787 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAenLysAspLysGln 806
Db 2730 GCATACAGAGAGTTGTGAATGAATAATGTTGCTGGACAGCACTGAAGAATAGAAA 2789
QY 807 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAenIleArgThrIleVal 826
Db 2790 GAACTAGAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACCTCCGAAACCGTTGTT 2849
QY 827 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 846
Db 2850 TCTTTGACTCAGGAGCAGAGAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATAC 2909
QY 847 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 866
Db 2910 AGAACTCTTTAGGAAGACACATCTTTGGAATATACATTTTCCCTTACCCAGGCAATG 2969
QY 867 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 886
Db 2970 ATGTATTTTCTATGCTGATGTTCCGGTTTGGAGCTTACTTGGTGGCATAAACTC 3029
QY 887 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 906
Db 3030 ATGAGCTTTGAGGATGTTCTGTAGTATTTTCAGCTGTTGCTTGGTGGCATGGCCGTG 3089
QY 907 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 926
Db 3090 GGGCAAGTCAGTTCAATTTGCTCTGACTATGCCAAGCCAAATATCAGAGCCCATC 3149
QY 927 PheAlaLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 946
Db 3150 ATCATGATCATTCGAAAAACCCCTTTGATTGACAGCTACAGCAGGAGGCCCTAATGCCG 3209
QY 947 AspThrCysGluGlyAanLeuGluPheArgGluValSerPhePheTyrProCysArgPro 966
Db 3210 AACACATTTGGAAGGAATGTACATTTGGTGAAGTTGTATTCAACTATATCCACCCGACCG 3269
QY 967 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 986
Db 3270 GATATCCAGTGTTCAGGACCTTGAGCTGGAGGTGAGAGGCCAGACGCTGGCTCTG 3329
QY 987 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 1006
Db 3330 GTGGGACAGCAGTGGCTGTGGGAAGACACAGTGTCTCCAGCTCTGGAGCGGTTCTACGAC 3389
QY 1007 ProValGlnGlyGlnValIlePheAspGlyValAspAlaLysGluLeuAanValGlnTyr 1026
Db 3390 CCTTGGCAGGGAAGTGTCTGTATGGCAAGAAATAAAGCGACTGAATGTTCAGTGG 3449
QY 1027 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 1046
Db 3450 CTCGAGCACACCTGGGCATCGTGTCCAGAGGCCATCTCTGTTGACTGCACATTTGCT 3509
QY 1047 GluAanIleAlaTyrGlyAspAenSerArgValValProLeuAspGluIleLysGluAla 1066
Db 3510 GAGAAATTCCTATGAGACAAACAGCCGGGTGTGTACAGGAAGAGATCGTGGAGGCA 3569
QY 1067 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAnThrGln 1086
Db 3570 GCAAGAGGCCCAACATACATGCTTCACTGCTGCTTAAATAATAGCACTAAA 3629

QY 1087 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 1106
Db 3630 GTAGGAGACAAAGGAACCTCAGCTCTCTGTGGCCAGAAACAGCATTGCCATAGCTCGT 3689
QY 1107 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 1126
Db 3690 GCCCTTTGTAGACAGCTCATATTTTGTCTTTGGATCAAGCCACGTCAGCTCGGATACA 3749
QY 1127 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1146
Db 3750 GAAAGTGAAAGGTTTCCAAAGAACCCCTGGCAAGCCAGAGAGCCGACCTGCATT 3809
QY 1147 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1166
Db 3810 GTGATTGCTCACCGCTCTCCACCATCCAGATGAGACTTAATAGTGTGTTTCAGAAAT 3869
QY 1167 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1186
Db 3870 GGCAGAGTCAAGGAGCATGGCAGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTT 3929
QY 1187 LysLeuValAsnAlaGlnSer 1193
Db 3930 TCAATGTCAGTGTCCAGGCT 3950
RESULT 14
AAZ94738
ID AAZ94738 standard; cDNA; 4646 BP.
XX
AC AAZ94738;
XX
DT 01-AUG-2000 (first entry)
XX
DE Human ATP binding cassette ABCB1 (MDR1) cDNA.
XX
KW ABCB1; ATP binding cassette; human; cholesterol; lipid disorder;
KW atherosclerosis; lipid disorder; dyslipidemia; psoriasis;
KW lupus erythematosus; diagnosis; gene therapy; MDR1;
KW multidrug resistance; chromosome 7q21, ss.
XX
OS Homo sapiens.
XX
FN WO200018912-A2.
XX
PD 06-APR-2000.
XX
PF 21-SEP-1999; 99WO-EP06991.
XX
PR 25-SEP-1998; 98US-0101706.
XX
PA (FARB) BAYER AG.
XX
PI Schmitz G, Klucken J;
XX
XX WPI; 2000-293151/25.
XX
DR Adenosine triphosphate binding proteins useful for identifying agents
PT for treating atherosclerosis and other inflammatory disorders -
XX
PS Claim 9; Page 110-112; 154pp; English.
CC
CC The present sequence is that of human ATP binding cassette
CC subfamily B protein ABCB1 cDNA. The cDNA was identified using a
CC differential display method in which monocytes from peripheral
CC blood were subjected to macrophage differentiation and cholesterol
CC loading with acetylated low density lipoproteins and subsequent
CC degrading with high density lipoprotein (HDL3) to identify
CC cholesterol sensitive genes. The gene maps to chromosome 7q21
CC and is also termed MDR1 (multidrug resistance). The invention
CC provides cholesterol-sensitive ABC genes (see AAZ94734-63). These
CC genes, and polypeptides encoded by them, can be used for diagnostic
CC and therapeutic applications, and for biochemical or cell-based
CC assays to screen for pharmacologically active modulator compounds

CC useful for the treatment of lipid disorders, atherosclerosis or
CC other inflammatory diseases such as psoriasis and lupus
CC erythematosus.

XX Sequence 4646 BP; 1371 A; 892 C; 1129 G; 1254 T; 0 other;

Alignment Scores:

Pred. No.:	0	Length:	4646
Score:	3545.00	Matches:	681
Percent Similarity:	75.55%	Conservative:	246
Best Local Similarity:	55.50%	Mismatches:	264
Query Match:	58.97%	Indels:	36
DB:	21	Gaps:	9

US-09-873-409-6 (1-1195) x AAZ9738 (1-4646)

Qy	1	MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu	20
Db	575	ATGGTGGTGGGAACCTTGGCTGCCATCATCCATGGGGCTGGACTTCTCTCATGATGCTG	634
Qy	21	ValLeuGlyGluMetSerAsp-----AsnLeuIleSer	31
Db	635	GTGTTTGGAGAAATGACAGATATCTTTGCAATGCGAGAAATTTAGAAAGATCTGATGTCA	694
Qy	32	GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe-----	43
Db	695	AACATCATCTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC	754
Qy	44	-----ArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr	61
Db	755	ATGACAGATGATGCTTATTATACAGTGAATTTGGTCTGGGGTGTCTGCTGCTTAC	814
Qy	62	IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln	81
Db	815	ATTGAGTTTTCATTTTGGTGGCTGGCAGCTGGAAGCAATAATACACAAATTAGAAACAG	874
Qy	82	PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu	101
Db	875	TTTTTTCATGCTATTAATGCGACGAGATAGGCTGGTTTGTGTGTCACCATGTTGGGGAG	934
Qy	102	LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle	120
Db	935	CTTAACACCCGACTTACAGATGATGCTCTTAAGATTATGAAGTATTGTGTCACAAAT	994
Qy	121	AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValIys	140
Db	995	GGAATGTTCTTTTCAGTCATGCAACATTTTTCACCTGGGTTTATAGTAGATTACAGCT	1054
Qy	141	GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla	160
Db	1055	GCTTGGAAAGCTAAACCTTGTGATTTTGGCCATCAGTCTCTTCTTGGACTGTGAGCTGCT	1114
Qy	161	AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla	180
Db	1115	GTCTGGGCAAGATACTATCTTCTTACTGATAAAGAACTCTTAGCGTATGCCAAAGCT	1174
Qy	181	GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln	200
Db	1175	GGAGCAGTAGCTGAAGAGCTCTTGGCAGCAATTAGACTGTGATGCAATTGGAGACAA	1234
Qy	201	GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys	220
Db	1235	AAAGAAGAACTTGAAGGTACAAACAAATTTTAGAAGAGCTAAAGAATTTGGGATAAG	1294
Qy	221	ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr	240
Db	1295	AAAGCTATTACGCCAATATTCTATAGTGTGCTTCTCTCTGATCATGCATCTTAT	1354
Qy	241	GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle	260
Db	1355	GCTCTGGCCCTCTGGTATGGGACCACTTGGTCTCTCAGGGGAA-----TATTCTATT	1408
Qy	261	GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla	280

Db	1409	GGCAAGTACTCACCTGTATCTTTCTGTATTAAATTTGGGGCTTTTAGTGTGGACAGGCA	1468
Qy	281	ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal	300
Db	1469	TCTCCAAGCATTTGAAGCATTTGCAATTCGAAGAGGAGCAGCTTATGAAATCTTCAAGATA	1528
Qy	301	IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle	320
Db	1529	ATTGATAATAAGCCAAGTATTGACAGCTATTTCGAAGAGTGGGCAAAACAGATAATATT	1588
Qy	321	GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys	340
Db	1589	AAGGGAATTTTGAATTCAGAAATGTTCACTTCAGTTACCCCATCTCGAAAAGAGTTAAG	1648
Qy	341	IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu	360
Db	1649	ATCTTGAAGGGCTGAACCTGAAGGTGAGAGTGGGACAGCGTGGCCCTGTTGGAAC	1708
Qy	361	AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp	380
Db	1709	AGTGGCTGTGGGAAGACACACAGCTGCAGCTGATGCAGAGCTCTATGACCCACAGAG	1768
Qy	381	GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp	400
Db	1769	GGGATGGTTCAGTGTGTGAGGACAGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA	1828
Qy	401	HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle	420
Db	1829	ATCATTTGGTGTGTGAGTGCAGGAACTGTATTGTTGGCCACCATAGTAGTCAAAACATT	1888
Qy	421	LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn	440
Db	1889	CGCTATGCGCGTGAATAATGTCAACATGATGATGATTGAGAAAGCTGTCAAGGAAGCAA	1948
Qy	441	AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly	460
Db	1949	GCCTATGACTTTTATCATGAACCTGCTCATAAATTTGACACCTCGTGTGGAGAGAGGG	2008
Qy	461	AlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn	480
Db	2009	GCCAGTTGAGTGTGGGAGAGGATCGCATTTGCACGTGCTGCTGTTGCCAAC	2068
Qy	481	ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla	500
Db	2069	CCCAAGATCTCTCTGCTGGATGAGGCCACGTGAGCTTGGACACAGAAAGCAAGCAGTG	2128
Qy	501	ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg	520
Db	2129	GTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTGCGACCCACCATTTGTGATAGCTCAT	2188
Qy	521	LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu	540
Db	2189	TGTCTACAGTTCTGTAATGCTGACGTCAATCGCTGCTGCTGCTGATGATGGAGTCAIT	2248
Qy	541	LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer	560
Db	2249	AAAGGAATCATGATGAATCATGAAGAGAAAGGCAATTTACTTCAAACTTGTCACAATG	2308
Qy	561	Gln-----AspIleLysLysAlaAspGluGlnMetGluSerMetThr	574
Db	2309	CAGACAGCAGGAAATGAAGTTGAATTAGAAATGACAGCTGATGAATCCAAAAGTGAAT	2368
Qy	575	TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal-----	589
Db	2369	GATGCTTGGAAATGCTTCAAAATGATTCAGATCCAGATCCAGTCTTAATAAGAAAAGATCA	2428
Qy	590	---LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys	607
Db	2429	CGTAGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACCAGAGGCTCTG	2488
Qy	608	GluIleSerLeuProGluValSerLeuLysIleLeuLysLeuAsnLysProGluTrp	627

Db 2489 GATGAAAGTATACCTCCAGTTTCCTTTGGAGGATTTAAGAGCTAAATTTAACTGAATG 2548
Qy 628 PropheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 647
Db 2549 CCTATTATTTGCTGGTGATTTTGTGCAATTTATAATGGAGGCTGCAACACGCAATTT 2608
Qy 648 SerIlePheAlaLysIleLeuThrMetPheGlyAsn---AsnAspLysThrThrLeu 666
Db 2609 GCATATATATTTCAAGATTTATAGGGTTTTCACGAATTTGATGATCCTGAAACAAA 2668
Qy 667 LysHisAspAlaGluLeuSerMetIlePheValLeuGlyValIleCysPheVal 686
Db 2669 CGACAGATAGTAATGTTTCTACTATTGTTCTAGCCCTTGAATATTTCTTTTAT 2728
Qy 687 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluLeuThrMetArgLeu 706
Db 2729 ACATTTTCTCCAGGGTTTCACATTTGGCAAGCTGGAGAGATCTCCAAAGCGGCTC 2788
Qy 707 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaIleAlaIlePheAspGluLysGlu 726
Db 2789 CGATACATGTTTCCGATCCATGCTCAGACAGATGTGAGTTGGTTGATGACCTTAA 2848
Qy 727 AsnSerThrGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 746
Db 2849 AACACCACTGGAGCATGACTACAGGCTCGCAATGATGCTGCTCAAGTTAAAGGGCT 2908
Qy 747 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 766
Db 2909 ATAGGTTCCAGGCTGCTGTAATTTACCAGATATAGCAATCTTGGGACAGGAATAT 2968
Qy 767 IleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeu 786
Db 2969 ATATCTTCTATCTATGTTGGCACTTAACACTGTACTCTTAGCAATGTACCCATCAT 3028
Qy 787 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 806
Db 3029 GCAATAGCAGGAGTTGTTGAAATGTAATGTTGCTGCAAGCACTGAAAGATAAGAAA 3088
Qy 807 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 826
Db 3089 GAACATAGAGGTCGTGGGAAGATCGCTACTGGAAGCAATAGAAACTTCCGAACCGTTGT 3148
Qy 827 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 846
Db 3149 TCITTGACTCAGGACAGAGATTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATAC 3208
Qy 847 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 866
Db 3209 AGAACTCTTTAGGAAAGCACACATCTTTGGAATTAATTTCTTCCATCCAGGCAATG 3268
Qy 867 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 886
Db 3269 ATGTATTTTCTATGCTGATGTTCCGGTTTGGAGCCCTACTTGTGGCACATAAACTC 3328
Qy 887 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 906
Db 3329 ATGAGCTTGAGGATGTTCTGTAGTATTTTACGCTGTTGCTGTGGTGCATGGCCGTG 3388
Qy 907 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 926
Db 3389 GGGCAAGTCAGTTTCATTTGCTCTGACTATGTCGCAAGCCAAATATCAGCAGCCACATC 3448
Qy 927 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 946
Db 3449 ATCATGATCATTTGAAAAACCCCTTTGATTTGACAGCTACAGCAGGAAGGCTTAATCCG 3508
Qy 947 AspThrCysGlyAsnLeuLeuPheArgGluValSerPheThrProCysArgPro 966
Db 3509 AACACATTTGGAGGAATGTCACATTTGGTGAAGTTGATTTCACTATCCACCCGACCG 3568
Qy 967 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 986
Db 3569 GACATCCAGTGCTTCAGGAGCTGAGGCTGGAGGTGAAGAGGCGCAGACGCTGCTG 3628

Qy 987 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 1006
Db 3629 GTGGGACAGTGGCTGTGGGAAGACACAGTGGTCCAGCTCTCTGGAGCGGTTCACGAC 3688
Qy 1007 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTyr 1026
Db 3689 CCCTTGGCAGGGAAGTGTCTGTATGTCGCAAAATAAAGACGACTGAATGTTCACTG 3748
Qy 1027 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 1046
Db 3749 CTCGAGACACCTGGCATCTGTCCTCCAGAGCCCATCTCTGTTCAGCTGCAGCATGCT 3808
Qy 1047 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 1066
Db 3809 GAGACATTTCCCTATGGACACAACAGCCGGTGGTGTCCACAGGAAGAGATCGTGAGGCA 3868
Qy 1067 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 1086
Db 3869 GCAAGGAGGCGCAACATACATCTCTGTCGTCGCAAAACAGCATTCCTTAATAATATAGCACTAA 3928
Qy 1087 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlnLysGlnArgLeuAlaIleAlaArg 1106
Db 3929 GTAGGAGACAAAGGAACCTCAGCTCTCTGTCGTCGCAAAACAGCATTCCTTAATAATATAGCACTAA 3988
Qy 1107 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 1126
Db 3989 GCCCTGTTAGACAGCTCATATTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACA 4048
Qy 1127 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1146
Db 4049 GAAAGTGAAGGTTTGTCCAAGAAAGCCCTGGCAAAAGCCAGAGAGCGGCACCTGCAT 4108
Qy 1147 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1166
Db 4109 GTGATTTGCTCACCGCTGTCACCATCCAGATTCAGACTTAATAGTGGTTCAGAAAT 4168
Qy 1167 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1186
Db 4169 GGCAGAGTCAGGAGCATGCGCAGCATCAGAGCTCTGCGACAGAAGGCATCTATTT 4228
Qy 1187 LysLeuValAsnAlaGlnSer 1193
Db 4229 TCAATGTCAGTGTCCAGGCT 4249
RESULT 15
AAD38994
ID AAD38994 standard; DNA; 4646 BP.
XX
AC AAD38994;
XX
DT 23-SEP-2002 (first entry)
DE Human mdrl gene.
XX
KW Human; haematologic malignancy; multidrug resistance; MDR; SUMO-1;
KW hypoxia inducible factor-1; small ubiquitin-like-modifier; HIP-1;
KW lymphoid disorder; chronic lymphoproliferative disorder; lymphoma;
KW myeloid disorder; lymphocytic leukaemia; thrombocythaemia; myeloma;
KW angogenic myeloid metaplasia; myeloid leukaemia; gene therapy;
KW polycythaemia vera; hypoxia responsive element; HRE; gene; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 425..4267
FT /tag= a
FT /product= "Human MDR protein"
XX
PN WC020234291-A2.
PD
XX 02-MAY-2002.

PF 25-OCT-2001; 2001WO-US49856.

PR 26-OCT-2000; 2000US-243542P.

PA (BGHM) BRIGHAM & WOMENS HOSPITAL INC.

XX Colgan SP;

XX WPI; 2002-471427/50.

XX P-PSDB; AAE24211.

DR Treating a subject (at risk of) having a hematologic malignancy or

XX multidrug resistance, e.g. lymphoma or myeloma, by administering

PT hypoxia inducible factor 1 binding molecules or small

PT ubiquitin-like-modifier-1 binding molecules -

XX Disclosure; Page 53-55; 92pp; English.

PS The invention relates to a method of treating a subject having or at

XX risk of developing a haematologic malignancy or multidrug resistance

CC (MDR). The method involves administering hypoxia inducible factor-1

CC (HIF-1) binding molecules or small ubiquitin-like-modifier (SUMO)-1

CC binding molecules or HIF-1-SUMO-1 complex modulators. mdrl-hypoxia

CC responsive element (HRE) binding molecules or antisense nucleic

CC acid molecules and SUMO-1 binding molecules or antisense molecules

CC are useful for treating a subject having or at risk of developing

CC haematologic malignancy or MDR (e.g. a lymphoid disorder or a myeloid

CC disorder). The lymphoid disorders include lymphocytic leukaemia or

CC chronic lymphoproliferative disorders e.g. lymphoma, myeloma or chronic

CC lymphoid leukaemia. The myeloid disorders include chronic or acute

CC myeloid leukaemia, e.g. angioleukemia myeloid metaplasia, essential

CC thrombocythaemia or polycythaemia vera. The invention is used in gene

CC therapy. The present sequence is human mdrl gene.

XX SQ Sequence 4646 BP; 1371 A; 892 C; 1129 G; 1254 T; 0 other;

Alignment Scores:

Pred. No.:	0	Length:	4646
Score:	3545.00	Matches:	681
Percent Similarity:	75.5%	Conservative:	246
Best Local Similarity:	55.50%	Mismatches:	264
Query Match:	58.97%	Indels:	36
DB:	24	Gaps:	9

US-09-873-409-6 (1-1195) x AAD38994 (1-4646)

Qy 1 MetIleLeuGlyIleleuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20

Db 575 ATGGTGGTGGGAACCTTGGCTGCCATCATCCATGGGGCTGGACCTTCCTCATGATGCTG 634

Qy 21 ValLeuGlyGluMetSerAsp-----AsnLeulleSer 31

Db 635 GTGTTGGAGAAATGACAGATATCTTGGCAATGTCAGGAAATTTAGAAATCTGATGTC 694

Qy 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe-----43

Db 695 AACATCACTAATAGAAAGTATCATCATGATCATACAGGGTCTTTCATGAATCTGGAGGAAC 754

Qy 44 -----ArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeullePheGlyTyr 51

Db 755 ATGACCATGATGCTGCTATTTATACAGTGAATTTGGTGGGGTGGTGGTGGTGGTGGTGG 814

Qy 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrIleArgIleArgIleGln 81

Db 815 ATTCAGTTTCATTTTGGTGGCTGGCAGCTGGAGAACAAATATACACAAATTTAGAAACAG 874

Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101

Db 875 TTTTTCATGCTATATGCGACAGGAGATAGCTGGTTGATGTCGACCATGCTGGGGAG 934

Qy 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspIle 120

Db 935 CTTAAACCCCGACTTACAGATGATGCTCTTAAGATTAATGAAGTTATTGGTGACAAAATT 994

Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140

Db 995 GGAATGTTCTTTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGGATTTACACGT 1054

Qy 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160

Db 1055 GGTGGAAAGCTAACCTTGTGATTTTGGCCATCACTCTCTTCTTGGAGCTGTCAGCTGCT 1114

Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180

Db 1115 GTCCTGGCAAGATACCTATCTTCACTTATGATAAAGAACTCTTAGCTATGCAAAAGCT 1174

Qy 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200

Db 1175 GCGAGCAGTAGCTGAAGAGCTCTTGGCAGCAATTAGAACTGTGATTGCAATTGGAGACAA 1234

Qy 201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220

Db 1235 AAGAAGAAGCTTTGAAAGGTACAAACAAAATTTAGAAAGCTTAAAGAAATTTGGGATAAG 1294

Qy 221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240

Db 1295 AAAGCTATTACAGCCAAATATTTCTATAGTGTCTTCTTCTCTGCTGATCTATCATCTTAT 1354

Qy 241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260

Db 1355 GCTCTGGCTCTTGGTATGGGACCACCTTGGTCTCTCAGGGGAA-----TATTCTATT 1408

Qy 261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280

Db 1409 GGCAAGTACTCACTGTATCTCTTTCTGTATTAAATTTGGGGCTTTTAGTGTGGCAGGCA 1468

Qy 281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300

Db 1469 TCTCAAGCATTTGAAGCATTTTCCAATTCAGAGGAGCGCTTATGAAATCTTCAGATA 1528

Qy 301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320

Db 1529 ATTGATAATAAGCCAAAGTATTTCAGACGCTATTTCGAAGAGTGGGCGACAAACAGATAATT 1588

Qy 321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340

Db 1589 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAAGTTAA 1648

Qy 341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360

Db 1649 ATCTTGAAGGGCTGAACTGAACTGAAGTGCAGAGTGGGAGCGGCTGCTGGTTGGAAC 1708

Qy 361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380

Db 1709 AGTGGCTGTGGGAAGAGCACACAGCTCCAGCTGATGTCAGAGGCTCTATGACCCACAGAG 1768

Qy 381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400

Db 1769 GGGATGGTGCAGTGTTCATGGAGCAGGATATTAGGACCAATAAATGTAAAGGTTTCTACGG 1828

Qy 401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420

Db 1829 ATCATTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1888

Qy 421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440

Db 1889 CGCTATGGCGTGAATAATGTCCATGGATGATGATGATGATGATGATGATGATGATGATGAT 1948

Qy 441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGly 460

Db 1949 GCCTATGACTTTTATCATGAACTGCTCATAAATTTGCACCCCTGTTGGAGAGAGGG 2008

Qy 461 AlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480

Db 2009 GCCCAGTTGAGTGGTGGGCGAGAGGATCGCCATTGGCAGCTGCCCTGGTTCGCAAC 2068

Db 4229 TCNATGTCAGTGTCCAGGCT 4249

Search completed: March 30, 2003, 03:20:36
Job time : 1405.91 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 30, 2003, 03:22:44 ; Search time 8219.42 Seconds
(without alignments)
2354.618 Million cell updates/sec

Title: US-09-873-409-6
Perfect score: 6012
Sequence: 1 MILGILASLVANGACPLMPL.....QELLNRDIYFKLVNAQSVQ 1195

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DEV=xlh
-MODE=frame_p2n.model -DEV=xlh
-Q/cgn2_1/USPTO.spool/US09873409/runat.27032003.115420.19240/app.query.fasta.1.7544
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09873409 @CGN 1.1 30544 @runat.27032003.115420.19240 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estovl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hun:*
19: em_gss_huv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_fam:*
24: em_gss_mus:*
25: em_gss_othr:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1092	18.2	669	13	BM013981	BM013981 603639619
2	1051	17.5	1033	14	BM926413	BM926413 AGENCOURT
3	1008	16.8	943	9	AL520322	AL520322 AL520322
4	995	16.6	760	13	BM016204	BM016204 603642659
5	926.5	15.4	929	14	B0882401	B0882401 AGENCOURT
6	918	15.3	545	9	AL040762	AL040762 DKFP2434C
7	856.5	14.2	948	13	BM471690	BM471690 AGENCOURT
8	854	14.2	2676	11	AK014319	AK014319 Mus muscu
9	834	13.9	998	14	BM904842	BM904842 AGENCOURT
10	771	12.8	894	12	BF584668	BF584668 602098406
11	770.5	12.8	1019	12	BG248052	BG248052 602359987
12	753	12.5	726	12	BG293345	BG293345 602390738
13	752.5	12.5	871	17	AZ682350	AZ682350 ENTKB16TF
14	751	12.5	944	12	BF796582	BF796582 602258463
15	744.5	12.4	872	11	AK020318	AK020318 Mus muscu
16	743.5	12.4	886	17	BH139685	BH139685 ENTNA47TR
17	734.5	12.2	932	17	AZ670821	AZ670821 ENTUN69TF
18	728.5	12.1	947	17	AZ683753	AZ683753 ENTIL96TF
19	726.5	12.1	886	17	AZ540627	AZ540627 ENTBQ18TF
20	723.5	12.0	880	17	AZ687805	AZ687805 ENTUJ52TF
21	721.5	12.0	939	14	BQ720763	BQ720763 AGENCOURT
22	720.5	12.0	897	17	AZ541090	AZ541090 ENTDS67TR
23	719	12.0	780	12	BG587938	BG587938 EST489713
24	711.5	11.8	913	17	BH155700	BH155700 ENTRO54TR
25	709.5	11.8	834	17	AZ548312	AZ548312 ENTRO77TR
26	708.5	11.8	853	17	AZ679807	AZ679807 ENTHI67TR
27	705.5	11.7	823	17	AZ532602	AZ532602 ENTCR50TR
28	702.5	11.7	1813	11	U66688	U66688 Homo sapien
29	697.5	11.6	861	14	BQ717101	BQ717101 AGENCOURT
30	692.5	11.5	1341	11	AY108285	AY108285 Zea mays
31	690.5	11.5	899	17	BH154857	BH154857 ENTRQ47TF
32	690	11.5	835	13	BJ349604	BJ349604 BJ349604
33	689.5	11.5	891	17	AZ682250	AZ682250 ENTKT68TR
34	688	11.4	1123	11	AY108485	AY108485 Zea mays
35	687	11.4	715	14	BQ869512	BQ869512 QGD6G09.Y
36	684	11.4	820	13	BJ356540	BJ356540 BJ356540
37	683	11.4	600	13	BM486593	BM486593 pgm2n.pk0
38	682.5	11.4	919	17	AZ690701	AZ690701 ENTV68TR
39	680	11.3	926	14	BQ123477	BQ123477 EST609053
40	677.5	11.3	795	12	BF133560	BF133560 601900192
41	674.5	11.2	815	13	BI409057	BI409057 602961127
42	674	11.2	2564	11	AK020022	AK020022 Mus muscu
43	671.5	11.2	803	12	BG298756	BG298756 602396681
44	670	11.1	852	17	BH720383	BH720383 BOH2294TF
45	669	11.1	547	13	BJ030711	BJ030711 BJ030711

ALIGNMENTS

RESULT 1
BM013981
LOCUS 603639619F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:5415556 5',
DEFINITION mRNA sequence.
ACCESSION BM013981
VERSION BM013981.1 GI:16528335
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 669)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL
COMMENT

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTp
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM2060 Row: f column: 05
High quality sequence stop: 664.
Location/Qualifiers
1. 669
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5415556"
/clone_lib="NIH_MGC_87"
/tissue_types="mammary adenocarcinoma, cell line"
/note="Organ: breast; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.383 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH MGC Library."
203 a 122 c 162 g 182 t

FEATURES

source

BM926413 1033 bp mRNA linear EST 12-MAR-2002
AGENCOURT_6600788 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5764845
5', mRNA sequence.

ACCESSION BM926413

VERSION BM926413.1 GI:19376792

KEYWORDS EST.

SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1033)

AUTHORS NIH-MGC http://mgi.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgabbs-r@mail.nih.gov

Tissue Procurement: Life Technologies, Inc.

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM21818 row: o column: 22

High quality sequence stop: 3

High quality sequence stop: 686.

FEATURES

source

Location/Qualifiers

1. 1033

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5764845"

/clone_lib="NIH_MGC_114"

/lab_host="DH10B"

/note="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: EcoRV (destroyed); RNA source anonymous pool of 6
male brains, age range 23-27 yo. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.5 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 019. Note:
this is a NIH MGC Library."

289 a 212 c 218 g 312 t

2 others

BASE COUNT

ORIGIN

Alignment Scores:

Pred. No.: 4 68e-104 Length: 1033

Score: 1051.00 Matches: 221

Percent Similarity: 94.49% Conservative: 2

Best Local Similarity: 93.64% Mismatches: 13

Query Match: 17.48% Indels: 3

DB: 14 Gaps: 0

US-09-873-409-6 (1-1195) x BM926413 (1-1033)

QY 266 ValPheSerValIleHisSerTyrCysIleGlyAlaValProHisPheGlu 285

Db 180 GTTTCTTTAGTCTAATCCATAGCAGTTATTGATTGGAGCAGCAGTCCCTCATTGAA 239

QY 286 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysPro 305

Db 240 ACCTTCGCAATACCCGAGGAGCTGCTTTTCATATTTTCCAGGTTCATTGATAAGAAACC 299

QY 306 SerIleAspAsnPheserThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 325
 Db 300 AGTATAGATAACCTTTCCACAGCTGATATAAACCTGAAATCCATAGAGAACTGTGGAA 359
 QY 326 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLysGlyLeu 345
 Db 360 TTTAAAAATGTTCTTTCAATATCCATCAAGACCATCTATCAAGATCTTGAAGGTCGTG 419
 QY 346 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 365
 Db 420 NATCTCAGATTAAGTCTGGAGAGACAGTCGCTTGGTCTCAATGCGAGTGGAG 479
 QY 366 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheLeuMetVal 385
 Db 480 AGTACGGTAGTCCAGCTTCTCGAGAGGTTATATGATCCGATGATGCTTATCATGGTG 539
 QY 386 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal 405
 Db 540 GATGAGATGACATCAGAGCTTTAAATGTGCGGCATTTATCGAGACCATATGAGGTGTT 599
 QY 406 SerGlnGluProValLeuPheGlyThrIleSerAsnAsnIleLysTyrGlyArgAsp 425
 Db 600 AGTCAAGAGCTGTTTGTTCGGGACCATCATCAGTAAACAATATCAAGTATGACGAGAT 659
 QY 426 AspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIle 445
 Db 660 GATGTGACTGATGAGAGATGAGAGAGCAGCAAGGAAAGCAAAATCGGTATGATTTATC 719
 QY 446 MetGluPheProAsnLysPheAsnThrLeuValGlyLysGlyAlaGlnMetSerGly 465
 Db 720 ATGGAGTTTCCCTAATAA-TTTAATACATTTGGTAGGGGANAAGGAGCTCCAATGAGTGA 778
 QY 466 GlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIle 485
 Db 779 ACCGAGAACCG-AGGATCCCAATGCTCTCTGCTTGTAGTTGGAACCCAGGAT-CTGAAT 836
 QY 486 LeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaVal 501
 Db 837 TTAGATGAAGCTAGCTGCTGCCCTGGATTGAGAAACCAAGTCAACTGCTC 884

RESULT 3
 AL520322
 LOCUS
 DEFINITION AL520322 LTI_NFL004_NBC2 Homo sapiens cDNA clone CS0DB006YC15 5
 prime, mRNA sequence.
 ACCESSION AL520322
 VERSION AL520322.1 GI:12783815
 KEYWORDS EST.
 SOURCE human.
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 943)
 AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
 Location/Qualifiers
 1..943
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CS0DB006YC15"
 /clone_lib="LTI_NFL004_NBC2"
 /sex="male"
 /tissue_type="neuroblastoma cells"
 /lab_host="DH108"
 /note="Organ: Brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and

cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 253 a 225 c 253 g 211 t
 ORIGIN

Alignment Scores:
 Score: 2.08e-99 Length: 943
 Pred. No.: 1008.00 Matches: 194
 Percent Similarity: 80.40% Conservative: 48
 Best Local Similarity: 64.45% Mismatches: 58
 Query Match: 16.77% Indels: 1
 DB: 9 Gaps: 0

US-09-873-409-6 (1-1195) x AL520322 (1-943)

QY 894 IleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAla 913
 Db 5 TTATTATTTCAGCTGTGTCTTTGGTGCATGGCGTGGGCAAGTCAGTTTCATTGCT 64
 QY 914 ProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLys 933
 Db 65 CTGACTATGCCAAAGCCAAATATCAGCAGCCCATCATCATGATCATTTGAAAAAAC 124
 QY 934 ProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeu 953
 Db 125 CTTTGAITGACAGCTACAGCAGGAGGCCTAATGCCAACAACATTTGGAAGAAATGTC 184
 QY 954 GluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGly 973
 Db 185 ACATTGGTGAAGTTGTATTCAACTATCCACCAGCCGACATCCCATGCTTCAGGA 244
 QY 974 LeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGly 993
 Db 245 CTGAGCTCGAGGTGAAGAAAGGCGCAGACGCTGGTCTCTGGTGGCAGCAGTGGCTGG 304
 QY 994 LysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeu 1013
 Db 305 AAGACACAGTGGTCCAGCTCTCTGGAGCGGTTCTACGACCCCTTGGCAGGAAAGTGTG 364
 QY 1014 PheAspGlyValAlaLysGluLeuAsnValGlnThrLeuArgSerGlnIleAlaIle 1033
 Db 365 CTTGATGGCAAGAAATAAAGCAGCTAATGTTTCAGTGGCTCCGAGCACACCTGGGCATC 424
 QY 1034 ValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAsp 1053
 Db 425 GTGTCCAGAGGCCCATCTCTTTGACTGCAGCATTTGCTGAGAACATTCCTATGAGAGAC 484
 QY 1054 AsnSerArgValValProLeuAspGluIleLysGluAlaAsnAlaAlaAsnIleHis 1073
 Db 485 AACAGCCGGGTGGTGTCAAGAGAGAGATCGTGAGGCGCAGCAAGAGGCGCAACATACAT 544
 QY 1074 SerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLysGlyAlaGln 1093
 Db 545 GCCTTCATCAGTCACTGCTTAATAATATAGCACTAAAGTAGGAGACAAAGGAACCTCAG 604
 QY 1094 LeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLys 1113
 Db 605 CTCTCTGGTGGCCAGAAACAAGCATTTGCCATAGCTGCTGCTGCTGTTAGACAGCCTCAT 664
 QY 1114 IleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGln 1133
 Db 665 ATTTTGTCTTTTGGATGAAGCCACGCTCTGGATACAGAAAGTGAAGAGTTGTCCAA 724
 QY 1134 HisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSer 1153
 Db 725 GAAGCCCTGGCAAAAGCCAGAGAGCCGACCTGCTGATTTGATTTGCTCACCCTGCTTC 784
 QY 1154 AlaIleGlnAsnAlaAspLeuIleVal-ValLeuHisAsnGlyLysIleLysGluGlnG 1173


```

|||||
Db 785 ACCATCCAGAAATGCAGACTTAAATAGTGTGTTTTCAGAAATGGCAGAGTCAAGGACATGG 844
|||||
Qy 1173 yThrHisGlnLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSe 1193
|||||
Db 845 CACGCATCAGCAGCTGCTGGCAGAAAGGCAATCTATTTTCAATGGTGTCCAGGC 904
|||||
Qy 1193 r 1193
Db 905 T 905

RESULT 4
BM016204 760 bp mRNA linear EST 30-OCT-2001
LOCUS 603642659F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:5418615 5',
DEFINITION mRNA sequence.
ACCESSION BM016204
VERSION BM016204.1 GI:165330558
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 760)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12068 row: e column: 16
High quality sequence stop: 740.

FEATURES
source
1..760
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5418615"
/clone_lib="NIH_MGC_87"
/tissue_type="mammary adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: Breast; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.383 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

BASE COUNT 229 a 150 c 187 g 193 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 3..76e-98 Length: 760
Score: 995.00 Matches: 203
Percent Similarity: 98.07% Conservative: 0
Best Local Similarity: 98.07% Mismatches: 2
Query Match: 16.55% Indels: 2
DB: 13 Gaps: 0

US-09-873-409-6 (1-1195) x BM016204 (1-760)

Qy 303 LysLysProSerIleAspAsnPhSerThrAlaGlyTyrLysProGluSerIleGluGly 322
Db 3 AAGAAACCCAGTATAGATAACTTTTCCACAGCTGGATATAACCTGAATCCATAGAAGGA 62
Qy 323 ThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeu 342
Db 63 ACTGTGAATTTAAATAGTTCTTTCAATATATCCATCAAGACCATCTATCAAGATTCTG 122
Qy 343 LysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGly 362

```

```

Db 123 AAAGGTCGAATCTCAAGATTAAGTCTGGAGAGACAGTGCCTTGGTCGGTCTCAATGGC 182
Qy 363 SerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspGlyPhe 382
Db 183 AGTGGGAAGAGTAGGTAGTCCAGCTCTTCGAGAGGTTATATGATCCGATGATGGCTTT 242
Qy 383 IleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIle 402
Db 243 ATCATGGTGGATGAGAAATGACATCAGAGCTTTAAATGTGGCGCATTCATCAGACCATATN 302
Qy 403 GlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnIleLysTyr 422
Db 303 GGAGTGTAGTCAAGAGCGCTGTTTGTTCGGACCAACCATCATAGTAACAATATCAAGTAT 362
Qy 423 GlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyr 442
Db 363 GGACGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAAATGCGTAT 422
Qy 443 AspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGln 462
Db 423 GATTTTATCATGAGTTCCTAATAATTTAATACATTTGTTAGGGGAAAAGAGGCTCAA 482
Qy 463 MetSerGlyGlyGlnLysGlnArgIleAlaIle-AlaArgAlaLeuValArgAsnProLy 482
Db 483 ATCAGTGGAGGGCAGAAACAGAGGATCGCAATTTGGCTCGTGCCTTAGTTTCGAAACCCCAA 542
Qy 482 sIleLeuIleLeu-AspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValG 502
Db 543 GATTCGTGATTTTCAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGCAAGTCAGCTGTT 602
Qy 502 InAlaLeuGluLys 507
Db 603 AAGCTGCACCTGGAGAG 619

RESULT 5
BM082401
LOCUS BM082401
DEFINITION AGENCOURT_8627902 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:6291782
5', mRNA sequence.
ACCESSION BM082401
VERSION BM082401.1 GI:22274409
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 929)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2492 row: c column: 15
High quality sequence stop: 677.

FEATURES
source
1..929
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6291782"
/clone_lib="NIH_MGC_43"
/tissue_type="normal pigmented retinal epithelium"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: eye; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5',

```


adaptor: GGCACAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH-MGC Library. "

BASE COUNT 281 a 176 c 237 g 234 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 1.8e-90 Length: 929
Score: 926.50 Matches: 188
Percent Similarity: 79.30% Conservative: 61
Best Local Similarity: 59.81% Mismatches: 58
Query Match: 14.41% Indels: 7
DB: 14 Gaps: 2

US-09-873-409-6 (1-1195) x BQ882401 (1-929)

QY 176 AlaTyrSerLysAlaGlyAlaValAlaGluValLeuSerSerIleArgThr-Valll 195
Db 1 GCGTATGCAAAAGCTGGAGCAGTAGCTGAAGAGGCTTGGCAGCAATTAGAACTGGTGAT 60
QY 195 eAlaPheArgAlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLyl 215
Db 61 TGCATTGGAGGACAAAGAAAGAACTTGAAGGTACAAACAAAATTTAGAAAGACTAA 120
QY 215 sAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePh 235
Db 121 AAGATTGGGATTAAGAAAGCTATTACAGCAATATTTCTATAGGTGCTTCTCTGCT 180
QY 235 eMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGl 255
Db 181 GATCATGTCATCTTATGCTCTGGCCCTCTGGTATGGACACACCTTGGTCTCTCAGGGGA 240
QY 255 uProGlyTyrThrIleGlyThrValLeuAlaValPheSerValIleHisSerSerty 275
Db 241 A-----TATTCTATTGGACAAAGTACTACTGATTTCTTTCTGATTAATTTGGGGCTTT 294
QY 275 rCySileGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPh 295
Db 295 TAGTGTGGACAGCATCTCCAGCATTTAGCATTTGCAATGCAAGAGGAGCAGCTTA 354
QY 295 eHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTy 315
Db 355 TGAATACTTCAAGATAATTTGATAATAAGCAAGTATTGACAGCTATTGCAAGAGTGGCA 414
QY 315 rLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSe 335
Db 415 CAACACAGATAATATTAAAGGAAATTTGGAAATTCAGAAATGTTCACTTACCTTACCATC 474
QY 335 rArgProSerIleLysIleLysGlyLeuAsnLeuArgIleLysSerGlyGluThrVa 355
Db 475 TGAAGAAGATTAAAGATCTTGAAGGCTTGAACCTTGAAGGTGCAGAGTGGGACAGGT 534
QY 355 lAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLe 375
Db 535 GGCCCTGGTTGGAACACAGTGGCTGGGAGAGACACACAGTCCAGCTGATGACAGGCT 594
QY 375 uTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnVa 395
Db 595 CTATGACCCACACAGGGGATGCTAGTGTGTGATGACAGGATATTAGGACCAATAATGT 654
QY 395 lArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrTh 415
Db 655 AGGTTTCTACGGGAATCATTTGGTGTGGTGTAGTACGAACCTGTATGTTTGTCCACCAC 714
QY 415 rIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAl 435
Db 715 GATAGCTGAAAACATTTCCCTATGSCCGTGAATAATGTCACCATGATGAGATTGAGAAAGC 774
QY 435 aAlaArgGluAlaAsnAla-TyrAspPheIleMetGluPheProAsnLysPheAsnThrL 455
Db 775 TGTCCAGGAAGCAATGCGCTATGACTTTTATCATGAACCTGCGCTTCAATAATTTGACACCC 834

QY 455 euValGlyGluLysGlyAlaGlnMetSerGlyGly---GlnLysGlnArgIleAlaIleA 474
Db 835 TGGTTGGANAAGAGAGGGCCCAATTGAGTGGTGGGCGCAACAACCAAGGATCTCCCTTGG 894
QY 474 lArg-AlaLeuVal-ArgAsnProLysIleLeu 484
Db 895 CACGTGGCCCTGGTTCCCAACCCCAAGATCTC 928
RESULT 6
LOCUS AL040762 545 bp mRNA linear EST 29-FEB-2000
DEFINITION DKFZp343C1815_r1 434 (synonym: htes3) Homo sapiens cDNA clone
ACCESSION DKFZp343C1815_5', mRNA sequence.
VERSION AL040762
KEYWORDS AL040762.1 GI:5409708
SOURCE EST.
ORGANISM human.
REFERENCE 1 (bases 1 to 545)
AUTHORS Bloecker H., Boecher, M., Brandt, P., Mewes, H.W., Gassenhuber, J. and
Wiemann, S.
TITLE EST (Bloecker, et al.)
JOURNAL Unpublished (1999)
COMMENT Contact: Bloecker H
MIPS
Am Klopferplatz 18a D-82152 Martinsried, Germany
This is the 5' sequence of the clone insert
Clone from S. Wilmann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wilmann@dkfz-heidelberg.de;
Sequenced by GBF (National Research Centre for Biotechnology Ltd.,
Braunschweig/Germany) within the cDNA sequencing consortium of the
German Genome Project.
s1 sequence also available.
This clone (DKFZp343C1815) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES
source

1..545
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZp343C1815"
/clone_lib="434 (synonym: htes3)"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/note="vector: pSport1; Site_1: NotI; Site_2: SalI"

BASE COUNT 171 a 89 c 138 g 147 t
ORIGIN

Alignment Scores:
Pred. No.: 5.84e-90 Length: 545
Score: 918.00 Matches: 181
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 15.27% Indels: 0
DB: 9 Gaps: 0

US-09-873-409-6 (1-1195) x AL040762 (1-545)

QY 306 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 325
Db 3 AGTATAGATAACTTTTCCACAGCTGGATATAAACCCTGATCATCAAGATCTCAAGAGGCTCG 62
QY 326 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu 345
Db 63 TTTTAAATGTTTCTTCAATATCATCAAGACCATCTATCAAGATCTCAAGAGGCTCG 122
QY 346 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 365
Db 123 AATCTCAGAAATTAAGTCTGGAGAGACAGTCGCGCTTGGTGGTCTCAATGGCAGTGGGAAG 182


```

QY 366 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 385
Db 183 AGTACGGTAGTCAGCTTCTGCAGAGGTTATATGATCCGATGATGGCTTTATCATGGTG 242
QY 386 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal 405
Db 243 GATCAGATACATCAGAGCTTTAAATGTGGGCATTTATCGAGCCATATTCGAGTGGTT 302
QY 406 SerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAsp 425
Db 303 AGTCAAGAGCGCTGTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGACGAGAT 362
QY 426 AspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIle 445
Db 363 GATGTGACTGATGAAGAGATGGAGAGACGACGAGGAGGAGCAATGCGTATGATTTATC 422
QY 446 MetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGly 465
Db 423 ATGGAGTTTCTTAATAATTTAATACATTGGTAGGGGAAAAAGAGCTCAATGAGTGA 482
QY 466 GlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIle 485
Db 483 GGGCAGAAACAGAGGATCGCAATTCGTGCTGCTTAGTTCGAAACCCCAAGATTCGATT 542
QY 486 Leu 486
Db 543 TTA 545

RESULT 7
LOCUS BM471690 948 bp mRNA linear EST 05-FEB-2002
DEFINITION AGENCOURT_6465349 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5539117
5', mRNA sequence.
ACCESSION BM471690
VERSION BM471690.1 GI:18520732
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 948)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-x@mail.nih.gov
Tissue Procurement: ATCC/DCTD/PTP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12233 row: b column: 14
High quality sequence stop: 569.
Location/Qualifiers
1. 948
/organism="Homo sapiens"
/db xref="taxon:9606"
/clone="IMAGE:5539117"
/clone_lib="NIH_MGC_72"
/tissue_type="melanotic melanoma"
/lab host="DHI10 (phage-resistant)"
/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr.
Average insert size 2 kb. Library constructed by Life
Technologies."
BASE COUNT 285 a 187 c 194 g 282 t
ORIGIN
Alignment Scores: 9.03e-83 Length: 948
Pred. No.:

```

```

Score: 856.50 Matches: 186
Percent Similarity: 89.67% Conservative: 5
Best Local Similarity: 87.32% Mismatches: 14
Query Match: 14.25% Indels: 8
DB: 13 Gaps: 1

US-09-873-409-6 (1-1195) x BM471690 (1-948)
QY 266 ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaValProHisPheGlu 285
Db 205 GTTTCCTTTAGTGTATATCATAGCAGTTATTCATTTGGAGCAGCAGCTCCCTCACTTGA 264
QY 286 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPro 305
Db 265 ACCTTCGCATAGCCGAGGAGCTGCTTTTCATATTTTCCAGSTATTATTAAGAACC 324
QY 306 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 325
Db 325 AGTATAGATAACTTTTCCACAGCTGGATATAAACCCTGAATCCATAGAGGAACCTGTGGA 384
QY 326 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu 345
Db 385 TTTAAATAATGTTTCCTTCAATTATCCATCAAGACCATCTATCAAGATTCGAAAGGCTG 444
QY 346 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 365
Db 445 AATCTCAGAAATTAAGTCTGGAGAGACAGTCGCTTGGTCTCAATGGCAGTGGAG 504
QY 366 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 385
Db 505 AGTACGGTAGTCAGCTTCTGCAGAGGTTATATGATCCGAGATGATGGCTTTATCATGGTG 564
QY 386 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIle-GlyValVal 405
Db 565 GATGAGAATGACATCAGAGCTTTAAATGTGGGCATATTCGAGACCATATTTGGAGTGGT 624
QY 405 lSerGlnGluPro--ValLeuPheGlyThrThrIleSerAsnAsnIleLysTyr-GlyAr 424
Db 625 TAGTCAAGAGCCCTGGTTTTCGGGACCACCATCAGTAACAATATCAAGTATGGGAGC 684
QY 424 AspAspValThrAspGluGlu-MetGluArgAlaAlaArgGluAla-AsnAlaTyrAsp 443
Db 685 AGATGATGGGACTGATGAACAAAATGGAAAAACCAACAAAGGGGAAATGA 744
QY 444 PheIleMetGlu-PheProAsnLysPheAsnThrLeuVal---GlyGluLysGlyAlaGl 462
Db 745 TTTATCATGAAATTTCCCAAAAAAATTTATCATTTGGAAGGGGAAAAAGGACCTTCA 804
QY 462 nMetSerGlyGlnLysGlnArg 470
Db 805 AGGAATGGAAGGGCCCAAAACCCCAAG 829

RESULT 8
LOCUS AK014319 2676 bp mRNA linear HTC 19-JAN-2002
DEFINITION Mus musculus 14, 17 days embryo head cDNA, RIKEN full-length
enriched library, clone:322401P09; ATP-binding cassette, sub-family
B (MDR/TAP), member 8, full insert sequence.
ACCESSION AK014319
VERSION AK014319.1 GI:12852089
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (strain: C57BL/6J) 14, 17 days embryo head cDNA to
mRNA, clone lib: RIKEN full-length enriched mouse cDNA library
Clone: 322401P09.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636

```



```
QY 35 ValGlnThrAsnThrTyrSerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyVal 54
|||
|||
|||
579 GTGCTGAGTCCCGTAAGCTCAGCGTCCAGCTGCTCTACTGTAC-----GGTGT 629
|||
|||
|||
QY 55 AlaAlaLeuIle---PheGlyTyrIleGlnIleSerLeuTyrIleIleThrAlaAlaArg 73
|||
|||
|||
630 CAGGAGCTGCTGACCTTGGATACCTAGTGTCTGTCTCCACATT-----GGTGAGCGC 683
|||
|||
|||
QY 74 GlnThrIleArgIleArgIleGlnPhePheHisSerValLeuAlaGlnAspIleGlyTyr 93
|||
|||
|||
684 ATGGCCATGACATCGGAAAGCCCTTTTCAGCTCCCTGCTCGGCAAGACATGCTTTC 743
|||
|||
|||
QY 94 PheAspSerCysAspIleGlyGluLeuAsnThrArgMet---ThrAspIleAspLys--- 111
|||
|||
|||
744 TTTGATGCCAAAAGACAGAGCGAGTAGTGTGCTGTGCTACTGTGTGCAAGATTC 803
|||
|||
|||
QY 112 -----IleSerAspGlyIleGlyAspLysIleAlaLeuLeuPhe 124
|||
|||
|||
804 AAGTCATCCTTCAAGCTTGTCTATCTCCAGGAGCTG----- 839
|||
|||
|||
QY 125 GlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu---ValLysGlyTyrLys 143
|||
|||
|||
840 ---CGCAGCTGCACCCAGGAGTGTGTAGCTGTGCTGTGCTGTGCTGTGCTGTGCT 896
|||
|||
|||
QY 144 LeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAlaCysSer 163
|||
|||
|||
897 CTTACCTGATGTGGCTGTGCTGCACACCCGCTCATGGAGTGGGCCACCTGTATGGGC 956
|||
|||
|||
QY 164 ArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaVal 183
|||
|||
|||
957 TCAGGCTCCGAAAGCTGTCTCGCCAGTGTCCAGGAGCATTTGCCAGGGCAACAGGTGA 1016
|||
|||
|||
QY 184 AlaGluGlnValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGlu 203
|||
|||
|||
1017 GCAGATGAGCCCTTGGCAATGTTCGAGCTGTGCGGCCCTTCGCGATGGAGAGAGGGAG 1076
|||
|||
|||
QY 204 LeuGlnArgTyrThrGlnAsnLeuLysAsp-----AlaLysAspPheGlyIle 219
|||
|||
|||
1077 GAGGAACGCTATCAGCAGAGACTGGAGTCATGCTGTGTAAGCAGAGAACTGGGC--- 1133
|||
|||
|||
QY 220 LysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhe-----PheMetAsn 237
|||
|||
|||
1134 ---AGGGGATCGGCTGTGTTCCAAAGGCTCTCCAAATCGCTTTCAACTGTATGGTCTTG 1190
|||
|||
|||
QY 238 GlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsnGlyGlu---Pro 256
|||
|||
|||
1191 GGCACC-----CTGTTTCATTTGGGGGCTCCCTTGTGGCTGCACAGCAGCTGAAA 1238
|||
|||
|||
QY 257 GlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerTyrCys 276
|||
|||
|||
1239 GGGGAGACCTCATGCTCTCTGCTGGTCCAGACAGTACAGAGTCTATGGCCAGC 1298
|||
|||
|||
QY 277 IleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGly-----AlaAla 294
|||
|||
|||
1299 CTCCTGTCCTGTTTGTGTGTCAG-----GTGGTACGTGGCTGAGTGCGGGA 1343
|||
|||
|||
QY 295 PheHisIlePheGlnValIleAspLysIlePheSerIleAspAsnPheSerThrAlaGly 314
|||
|||
|||
1344 GCCCGAGTCTTCGAATACATGCGCCCTGAGCCCTGTCTATCCCA-----TTGACCGGGGGC 1397
|||
|||
|||
QY 315 Tyr-----LysProGluSerIleGlyThrValGluPheLysAsnValSerPheAsn 332
|||
|||
|||
1398 TACTGCAATCCCAACAGGACATTTGTTGTTCCATCATCTTCCAAATGTACCTTCAGC 1457
|||
|||
|||
QY 333 TyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGly 352
|||
|||
|||
1458 TACCCCTGCAGACCTGGCTTCAATGTCTCAAGGACTTCCACCTGAAGCTGCCCTCTGCG 1517
|||
|||
|||
QY 353 GluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValGlnLeuLeu 372
|||
|||
|||
1518 AAGATTGTGGCTCTTGTGGGCCAGTCTGGGGAGGAGAACACACAGTGTGCTCTCTGCTG 1577
|||
|||
|||
```

```
QY 373 GlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAla 392
|||
|||
|||
1578 GAACGGCTTCTATGACCTTGAGCTCGTGGTACGCTTGGATGGCGCATGACCTCGGAAC 1637
|||
|||
|||
QY 393 LeuAsnValArgHisTyrArgAspHis---IleGlyValValSerGlnGluProValLeu 411
|||
|||
|||
1638 CTCACCCCTCTGCTCGGCGCCAGGTCTATAGTTTCATCAGCAGGAGCCAGTCTCTG 1697
|||
|||
|||
QY 412 PheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGlu 431
|||
|||
|||
1698 TTTGCAACACCATCATCGAGATATCCGATTGGGAAGCTGATGCTTCCGATGAAGAG 1757
|||
|||
|||
QY 432 MetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLys 451
|||
|||
|||
1758 GTGTACACAGCTGCACGAGAACCAATGCCAGATTCATCAGCAGCTTCCCGCATGGC 1817
|||
|||
|||
QY 452 PheAsnThrLeuValGlyLysGlyAlaGlnMetSerGlyGlnLysGlnArgIle 471
|||
|||
|||
1818 TACAGCAGCTGTGGTGGTGGCGGGGCAACCTTGTGTGGTGGCCAGAGCAGCGCTA 1877
|||
|||
|||
QY 472 AlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSer 491
|||
|||
|||
1878 GCCATCGACGTGCCCTCATCAGCAGCCACAGTCTGATCTCGACGAGGCCACCAT 1937
|||
|||
|||
QY 492 AlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGly 511
|||
|||
|||
1938 GCGCTAGATGCAGATCCGAGAGGGTGTACAGAGGCCCTCGACCGGCCAGCTGTGGC 1997
|||
|||
|||
QY 512 ArgThrThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleVal 531
|||
|||
|||
1998 CGCACCGTGTGGTATTGCCACCGGCTGTAGTCTGCTGCGTGCAGCCCATCATCAT 2057
|||
|||
|||
QY 532 ThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArg 551
|||
|||
|||
2058 GTCATGCCCAATGCCCAAGTCTGTGAGGCTGGGAGCCACAGAGACTCTCTTAAAAAGGC 2117
|||
|||
|||
QY 552 GlyLeuTyrTyrSerLeuValMetSerGlnAspIle 563
|||
|||
|||
2118 GGGCTTATTGAGAGCTTATCCGAGACAAAGCCCTG 2153
|||
|||
|||
RESULT 9
BM904842 998 bp mRNA linear EST 12-MAR-2002
LOCUS AGENCOURT_6699581 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5557655
DEFINITION 5', mRNA sequence.
ACCESSION BM904842
VERSION BM904842.1 GI:19355221
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 998)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DTF
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAMI2279 row: f column: 24
High quality sequence stop: 738.
FEATURES
Location/Qualifiers
1..998
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5557655"
/clone_lib="NIH_MGC_72"
source
```


[illegible]

QY 326 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLysLeuLysGlyLeu 345
 Db 184 TTTAAGATATTCATCTCAGTTACCATCTCGAAGAAGCTTCAGATCTTGAAGGCCCTC 243

QY 346 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 365
 Db 244 AATCTGAAGGTGAAGAGCGACAGCGTGGCTGGTGGCAACAGTGGCTGTGGAAGA 303

QY 366 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspSerGlyPheIleMetVal 385
 Db 304 AGCACAACTCTCAGCTGTGTCGAAGGCTCTACGACCCCTAGATGCGATGCGTATC 363

QY 386 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal 405
 Db 364 GACGGACAGACATCAGAACCATCATGTGAGGTATCTGAGGAGATCATGTGTGTGGT 423

QY 406 SerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAsp 425
 Db 424 AGTCAGAACCTGTGTCTGTTTGGCCACCAGCATCGCCGAGAACATTCGTATGCGCGAGAA 483

QY 426 AspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIle 445
 Db 484 GATGTCTACCATGGATGATGATGAGAAAGCTGTCAAGGAAGCCATGCTATGACTTCATC 543

QY 446 MetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGlyAlaGlnMetSerGly 465
 Db 544 ATGAACTGCCCCCAATTTTGACACCTGTTGGTGAGAGAGGGGCGCAGTGTGGG 603

QY 466 GlyGlnLysGlnArgIleAlaAlaArgAlaLeuValArgAsnProLysIleLeuIle 485
 Db 604 GGACAGAAACAGAGAAATCGC-ATTGCCCGGGCCCTGTCGCGCAATCCAAAGTGCCTTC- 661

QY 486 LeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaIle 505
 Db 662 GTGACGAAGCCACCTCAGC-CTGGATACAGAAAGTGAAGCTGGGGTTCAGGCGGCACCT 720

QY 505 uGluLysAlaSerLysGlyArgThrIle-----ValValAlaHisArgLe 521
 Db 721 AGATAGGTACAAAGCGCGGACCACTATGGGAGCTCATCGCTGTATCCGTCGAAGCT 780

QY 521 uSerThrIleArgSerAlaAspLeuIleVal-----ThrLeuLysAspGlyMetLeuAl 539
 Db 781 GACGCCCATCGCGCGGAAGGGGG-CTTGTGTGGAAACGACTCTGAAAGC-----TC 830

QY 539 aGluLysGlyAlaHis 544
 Db 831 TGAAGGAGGTACTCAT 846

RESULT 11
 LOCUS BG248052 1019 bp mRNA linear EST 13-FEB-2001
 DEFINITION 602359987F1 NCI_CGAP_Mam Mus musculus cDNA clone IMAGE:4488404 5', mRNA sequence.

ACCESSION BG248052
 VERSION BG248052.1 GI:12757867
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 1019)
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 TITLE NIH-MGC http://mgi.nci.nih.gov/
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
 COMMENT Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 EMAIL: cgapbs-rc@mail.nih.gov
 TISSUE: Tissue Procurement: Gilbert Smith, Ph.D.
 CDNA LIBRARY PREPARATION: Life Technologies, Inc.
 CDNA LIBRARY ARRAYED BY: The I.M.A.G.E. Consortium (LLNL)
 DNA SEQUENCING BY: Incyte Genomics, Inc.
 CLONE DISTRIBUTION: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM10334 row: n column: 21

High quality sequence stop: 650.

FEATURES

source

1..1019
 /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:4488404"
 /clone_lib="NCI_CGAP_Mam1"
 /tissue_type="tumor, biopsy sample"
 /dev_stage="10 months, virgin"
 /lab_host="DH10B"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
 Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
 Library constructed by Life Technologies. Investigator
 providing samples: Gilbert Smith, NIH"

BASE COUNT 276 a 262 c 283 g 196 t
 ORIGIN

Alignment Scores:

Pred. No.: 2,85e-73 Length: 1019
 Score: 770.50 Matches: 154
 Percent Similarity: 79.17% Conservative: 36
 Best Local Similarity: 64.17% Mismatches: 47
 Query Match: 12.82% Indels: 3
 DB: 12 Gaps: 1

US-09-873-409-6 (1-1195) x BG248052 (1-1019)

QY 950 GluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPhe 969

Db 4 GAAGGAATATGTCATTTAGTGGAGTCGTGTCAACATATCCACCAGCCAGCATCCCA 63

QY 970 IleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySer 989

Db 64 GTGCTTCAGGGGTGAGCCTTGAGGTGAAGAGGGCCAGCCGCTGGCTGGGGCAGC 123

QY 990 SerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGln 1009

Db 124 AGTGGCTGGGGAGAGCACAGTGTTCAGTGTCTCGAGGGCTTCTACGACCCCATGGCT 183

QY 1010 GlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSer 1029

Db 184 GGATCAGTGTCTCAGTGGCAAGAAATAAAGCAACTGAATGTCCAGTGGCTCCGAGCA 243

QY 1030 GlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIle 1049

Db 244 CAGCTGGGCATTGTGTCCCAAGAGCCCATCTCTTTGACTGCGAGCATCGCAGAGAACAT 303

QY 1050 AlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAla 1069

Db 304 GCCTACGGAGACACAGCCGGTGTGTCTTATGAGAGATTTGTGAGGGCAGCCAGGAG 363

QY 1070 AlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeu 1089

Db 364 GCCAATCCACCAGTTTCATCGACTCGCTACCTGATAAATAACAACACAGTAGGAGAC 423

QY 1090 LysGlyAlaGlnLeuSerGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeu 1109

Db 424 AAAGGCACCTCAGCTGTGGGTGGGAGAGCAGCGCATGCCATCGCAGCCCTCGTC 483

QY 1110 GlnLysProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspAsnSerGlu 1129

Db 484 AGACAGCTCACATTTTACTTCTGGACGAAGCAACATCAGCTCTGGATACAGAAAGTGA 543

QY 1130 LysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValThr 1149

Db 544 AAGGTTGTCCAGGAAGCGCTGGCAAAAGCAGGAAAGCGGACCTGTCATTGTGTATCGCT 603

QY 1150 HisArgLeuSerAlaIleGlnAsnAlaAspLeuIle-ValValLeuHis-AsnGlyLysI 1169

Db 604 CACCGCTGTCCACCATCCCAAGAACCGGAGCTTGTATCGGTGATTCAGAGAACCGGAGG 663

Qy	1169	leIys---GgUgInGlyThrHisGlnGluLeuLeuAraGAsnAraGspIleTyrPhe	1186
Db	664	TCRAAGGAGCAGCGGCACCCCAACAGCAAGTGGTGGCGCAGAGGGCATCTACTTC	719
RESULT 12			
BG293345			
LOCUS	602390738F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:4502552 5',	726 bp	linear EST 21-FEB-2001
DEFINITION	mRNA sequence.		
ACCESSION	BG293345		
VERSION	BG293345.1	GI:113052943	
KEYWORDS	EST.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	1 (bases 1 to 726)		
TITLE	NIH-MGC http://mgc.nci.nih.gov/.		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgapbs-x@mail.nih.gov		
	Tissue Procurement: The Cepko Laboratory		
	cDNA Library Preparation: Life Technologies, Inc.		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LLNL at:		
	Plate: LLAM10371 row: 1 column: 09		
	High quality sequence stop: 658.		
FEATURES	Location/Qualifiers		
source	1..726		
	/organism="Mus musculus"		
	/db_xref="taxon:10090"		
	/clone="IMAGE:4502552"		
	/clone_lib="NIH_MGC_94"		
	/tissue_type="retina"		
	/lab_hosts="DH10B (phage-resistant)"		
	/note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;		
	Site 2: SalI; Cloned unidirectionally; oligo-dr primed.		
	Average insert size 3.3 kb. Library enriched for		
	full-length clones and constructed by Life Technologies.		
	Note: this is a NIH_MGC Library."		
BASE COUNT	189 a 220 c 190 g 127 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	1..28e-71	Length:	726
Score:	753.00	Matches:	152
Percent Similarity:	84.29%	Conservative:	25
Best Local Similarity:	72.38%	Mismatches:	30
Query Match:	12.52%	Indels:	3
DB:	12	Gaps:	0
US-09-873-409-6 (1-1195) x BG293345 (1-726)			
Qy	984	ValAlaPheValGlySerSerGlyCysGlyLyssSerThrSerValGlnLeuLeuGlnArg	1003
Db	2	CTGGCCCTGGTGGGCAGCAGTGGCTGCGGGAGAGACACAGTGGTCCAGCTGCTCGAGGCG	61
Qy	1004	LeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLySGluLeuAsn	1023
Db	62	TTCTATGACCCCATGGCTGGATGATGCTCTTAGATGGTCAAGAGCAAGAAACTCAAT	121
Qy	1024	ValGlnTTrpLeuAraGSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCys	1043
Db	122	GTCCAGTGGCTCCGAGCTCAACTGGGCATTGTGTCCGAGGACCACTCTCTTTGACTGC	181
Qy	1044	SerTleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIle	1063
Db	182	AGCATCGCAGAGAATCATCGCTATGAGACAAACAGCCGGGTCTGTCCTCATGATCAGATT	241

[illegible]

H.O. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B. Borell, Oxford University Press, 1999).

BASE COUNT 349 a 117 c 174 g 231 t
ORIGIN

Alignment Scores:
Pred. No.: 2,03e-71 Length: 871
Score: 752.50 Matches: 151
Percent Similarity: 71.64% Conservative: 46
Best Local Similarity: 54.91% Mismatches: 75
Query Match: 12.52% Indels: 3
DB: 17 Gaps: 1

US-09-873-409-6 (1-1195) x AZ682350 (1-871)

Qy 290 AlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsn 309

Db 45 GCTAAAGTTCGCCGATTAATGTTTATCAACAATTCATAGATCCAGATATGATTTGT 104

Qy 310 PheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnVal 329

Db 105 CAGTCTATTGGAGGTGAATGTCCAAGTGAAGTGAATGGAATATTAGATTGAGGATGT 164

Qy 330 SerPheAsnTyrProSerArgProSerIleLysLysLysGlyLeuAsnLeuArgIle 349

Db 165 CAATTTGTTTATCCACAGAGCTGCTCATCATGTTATTAAGAGCTTGACCTTGAAT 224

Qy 350 LysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValVal 369

Db 225 AAGAAAGGACAAACAATTCGATTAGTTGGAGCAGCAGTGTGGAGTCAACTACTATT 284

Qy 370 GlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAsp 389

Db 285 CAATTAATCCAAAGAAATTAATGATCCAAATGGTGAAGAGTAACATTAGACGGAAAGAT 344

Qy 390 IleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluPro 409

Db 345 ATACGAGTGAATATCAATGTTAAGAAATCAATAGATTAGTTGGCAAGAACCA 404

Qy 410 ValLeuPheGlyThrIleSerAsnAsnIleLysTyrGlyArgAspAsp----- 426

Db 405 GTGTGTTTTCAGGAAACAATTCGAGAAATATTATGCTTCGAGCTAAAGAGGACCA 464

Qy 427 ValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMet 446

Db 465 CCAAGTGAAGAGAGATGATGTAATGTCTAAATATGCAATGCATGACATGCTTCTTCT 524

Qy 447 GluPheProAsnLysPheAsnThrLeuValGlyGlyLysGlyAlaGlnMetSerGlyGly 466

Db 525 AAATCTCAGAGGATATGACACAAATATGGAGAAAGAGGACATTTATTCAGGAGGA 584

Qy 467 GlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeu 486

Db 585 CAAAAACAAGAAATTCGAATTCGACGTGATTTGATTCGAAACCCATCTATTCTTCTTCT 644

Qy 487 AspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaLeuGlu 506

Db 645 GATCAAGCTACATCAGCACTTGATACCAAGAGTGAAGAGATTTGACAAAGACACTTGA 704

Qy 507 LysAlaSerLysGlyArgThrIleValValAlaHisArgLeuSerThrIleArgSer 526

Db 705 AAGCATCTAAGAGAGAACACAAATTTATTGTACACATGACATCAACCTGTAGAAT 764

Qy 527 AlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGlu 546

Db 765 GCAGATAAAATATGTGTTATTCATCAAGGAGAAATTTATTGAACAAGGAAACATCAAGAA 824

Qy 547 LeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln 561

Db 825 TTAATGATTTGAAGAGGAACATATTTATGATTTAGTCAAAAGACAA 869

RESULT 14
BF796582
LOCUS
DEFINITION

ACCESSION
VERSION
KEYWORDS
SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BF796582 944 bp mRNA linear EST 12-JAN-2001
602258463F1 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:4341710 5',
mRNA sequence.

BF796582 GI:12101636

EST.

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 944)

NIH-MGC http://mgs.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: c9apbs@mail.nih.gov

Tissue Procurement: Louis Staudt, M.D., Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM9955 row: n column: 15

High quality sequence start: 9

High quality sequence stop: 669.

Location/Qualifiers

1..944

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:4341710"

/tissue_type="lymphoma, cell line"

/lab_host="DH10B (phage-resistant)"

/note="Organ: lymph; Vector: pCMV-SPORT6; Site 1: NotI;

Site 2: SalI; Cloned unidirectionally; oligo-dT primed.

Average insert size 1.867 kb. Library enriched for

full-length clones and constructed by Life Technologies.

Note: this is a NIH_MGC Library."

BASE COUNT 265 a 241 c 237 g 201 t

ORIGIN

Alignment Scores:

Pred. No.: 3,43e-71 Length: 944

Score: 751.00 Matches: 166

Percent Similarity: 74.62% Conservative: 31

Best Local Similarity: 62.88% Mismatches: 60

Query Match: 12.49% Indels: 12 9

DB: 12 Gaps: 2

US-09-873-409-6 (1-1195) x BF796582 (1-944)

Qy 916 TyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsn 935

Db 39 TATGCTAAAGCTAAGCTGCTGCAGCCCACTTATTCATGCTTTCGAAAGACAACTCTG 98

Qy 936 IleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPhe 955

Db 99 ATTGACAGCTACAGTGAAGAGGGGCTGAAGCCCTGATAAATTTTGAAGGAAATATAACATTT 158

Qy 956 ArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSer 975

Db 159 AATGAGTCTGTTCAACTATCCACCCGAGCAACGTCAGTGTCTCAGGGGCTGAGC 218

Qy 976 LeuSerIleGluArgGlyLysThrValAlaPheValIcylSerSerGlyCysGlyLysSer 995

Db 219 CTGAGGTGAAGAAAGCCAGACACTAGCCCTGCTGGCAGCAGTGTCTGTGGGAGAGC 278

Qy 996 ThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAsp 1015

Db 279 ACGGTGCTCCAGCTCCTGGAGCGGTTCTACGACCCCTTGGCGGGGACAGTGTCTTCGAT 338


```
/strain="C57BL/6J"
/db_xref="FANTOM DB:9230106F14"
/db_xref="MGD:MGI:1911565"
/db_xref="taxon:10090"
/clone="9230106F14"
/sex="male"
/tissue_type="epididymis"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
86. .-872
/notes="data source:SPTR, source key:P21447, evidence:ISS
putative
similar to MULTIDRUG RESISTANCE PROTEIN 3 (P-GLYCOPROTEIN
3) (MDR1A)"
/codon_start=1
/protein_id="BAB32070.1"
/db_xref="GI:12860873"
/db_xref="MGD:MGI:1924956"
/translation="MANSRITGLQETNRYGPLEQVPKVGQNAVGPTEIFRFADNL
DIVLTLGLASMGATVPLMSVLGEISDHLINGCLVQTNRTKYQNCSTQEKLINE
DIIVLTVYIGIGAAALIFGYQISFWVITAAROTTRIRKOFHSLAODISWFDGSD
ICELNTRMTGDKLDCGIGDKIPLMFQNISGFSIGLIVSLIKSWKLSLVLSPLI
MASSALCSRLSIISLTKELDVAGAVAEALSSIQVTAFGA"
BASE COUNT 262 a 178 c 187 g 245 t
ORIGIN

Alignment Scores:
Pred. No.: 1.53e-70 Length: 872
Score: 744.50 Matches: 151
Percent Similarity: 80.37% Conservative: 21
Best Local Similarity: 70.56% Mismatches: 27
Query Match: 12.38% Indels: 15
DB: 11 Gaps: 2

US-09-873-409-6 (1-1195) x AK020318 (1-872)
Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 230 ATGACTCTGGGAATATTAGCATCCATGATAAATGAGCCCGCTCCTTTAAATGTCCTG 289
Qy 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
Db 290 GTTTTAGGAGAAATAAGTGATCATTTAATTAATGATGCGCTAGTACAAACTAACAGAACT 349
Qy 41 SerPhe-----PheArgLeuThr 46
Db 350 AAATATCAGAACTGTTCTCAGACTCAAGAAAGCTGAATGAAGATATCATTTGTTGACT 409
Qy 47 LeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeu 66
Db 410 CTATATTATATTGGAATAGGAGCAGCTGCCCTCATTTTGGCTATGTACAGATTTCTCTC 469
Qy 67 TrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGlnPhePheHisSerVal 86
Db 470 TGGGTCAATACTGCAGCCGCGCAACACACAGAATCCGAAAACAGTTTTCATTCAATT 529
Qy 87 LeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGluLeuAsnThrArgMet 106
Db 530 TTGGCAAGACATCAGCTGGTGTGGTGGCAGTGCATCTGTGAACCTTAACACCCCGCATG 589
Qy 107 Thr---AspIleAspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGln 125
Db 590 ACTGGTGACATCAACAACTCTGTGATGGTATTGGAGATAAGATCCCTCTGATGTTTCAG 649
Qy 126 AsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLysGlyTyrLysLeuThr 145
Db 650 AACATATCTGGGTTTCTATTGGCCTGGTGATAAGTTTGTAAAAAGCTGGAAACTCTCC 709
Qy 146 LeuValThrLeuSerThrProLeuIleMetAlaSerAlaAlaCysSerArgMet 165
Db 710 CTGGTGGTTCTGTCTACATCTCTCTCATATGGCTTCATCGGCACGTGCTCTAGGATG 769
Qy 166 ValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGlu 185
```

```
Db 770 ATTATCTCATTACCAGCAAGGAGCTGGATGCTATTCCAAAGCTGGGCTGTGGCTGAA 829
Qy 186 GluValLeuSerSerIleArgThrValIleAlaPheArgAla 199
Db 830 GAAGCCTTGTCTATCATCCAAACAGTCACAGCCTTTGGAGCC 871
```

Search completed: March 31, 2003, 13:51:24
Job time : 8245.55 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 31, 2003, 02:24:20 ; Search time 175.002 Seconds
(without alignments)
2094.145 Million cell updates/sec

Title: US-09-873-409-6

Perfect score: 6012

Sequence: 1 MILGILASLVNGACPLMPL.....QELLRRNDYFKLVNAQSVQ 1195

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09873409/runat_27032003.115422.19259/app_query.fasta_1.7544
-DB=Issued Patents NA -QPMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09873409 @CGN 1.1.578 @runat_27032003.115422.19259 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_WMAP -LARGQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:*

1: /cgn2_6/ptodata/2/ina/5A COMB.seq:*

2: /cgn2_6/ptodata/2/ina/5B COMB.seq:*

3: /cgn2_6/ptodata/2/ina/6A COMB.seq:*

4: /cgn2_6/ptodata/2/ina/6B COMB.seq:*

5: /cgn2_6/ptodata/2/ina/PCTUS COMB.seq:*

6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3545	59.0	4646	1	US-08-181-471-2
2	3545	59.0	6505	2	US-08-793-610-5
3	3545	59.0	9318	2	US-08-793-610-6
4	3543.5	58.9	4264	2	US-08-784-649A-1
5	3543.5	58.9	4264	2	US-08-784-649A-5
6	3543	58.9	4669	6	5206352-3
7	3527	58.7	4669	2	US-08-752-447-1
8	3527	58.7	4669	4	US-09-316-167-1
9	3517.5	58.5	4233	3	US-09-120-513-1
10	3517.5	58.5	4233	4	US-09-450-105-1
11	3497	58.2	4669	2	US-08-583-276-18
12	2282.5	38.0	4002	2	US-08-996-545-1

13	2282.5	38.0	4002	2	US-08-996-545-3	Sequence 3, Appli
14	2282.5	38.0	4002	4	US-09-328-320-1	Sequence 1, Appli
15	2282.5	38.0	4002	4	US-09-328-320-3	Sequence 3, Appli
16	2280	37.9	2726	1	US-08-461-823-1	Sequence 1, Appli
17	2273.5	37.8	4047	2	US-08-612-734B-1	Sequence 1, Appli
18	2211	36.8	4800	2	US-08-612-734B-3	Sequence 3, Appli
19	2166	36.0	3924	1	US-08-395-246C-1	Sequence 1, Appli
20	2116	35.2	4224	1	US-08-612-521-1	Sequence 1, Appli
21	1731.5	28.8	6143	1	US-08-612-521-3	Sequence 3, Appli
22	1703	28.3	3924	2	US-08-996-644-3	Sequence 3, Appli
23	1703	28.3	3924	3	US-09-352-552-3	Sequence 3, Appli
24	1703	28.3	3927	2	US-08-996-644-1	Sequence 1, Appli
25	1703	28.3	3927	3	US-09-352-552-1	Sequence 1, Appli
26	1664	27.6	3792	4	US-08-232-537-1	Sequence 1, Appli
27	1659.5	27.6	3792	4	US-09-351-224E-10	Sequence 10, Appli
28	1525	25.4	3999	4	US-09-351-224E-9	Sequence 9, Appli
29	1109	18.4	13188	4	US-08-961-527-70	Sequence 70, Appli
30	1057	17.6	7760	4	US-08-961-527-63	Sequence 63, Appli
31	996.5	16.6	441529	4	US-09-103-840A-1	Sequence 1, Appli
32	994.5	16.5	4403765	4	US-09-103-840A-2	Sequence 2, Appli
33	945.5	15.7	7186	4	US-08-961-527-39	Sequence 39, Appli
34	938	15.6	6492	4	US-08-961-527-188	Sequence 188, Appli
35	761.5	12.7	2376	1	US-08-394-880B-1	Sequence 1, Appli
36	718	11.9	1749	4	US-09-134-001C-1893	Sequence 1893, Appli
37	698	11.6	2061	4	US-09-061-764A-17	Sequence 17, Appli
38	687.5	11.4	5889	1	US-08-463-092B-5	Sequence 5, Appli
39	687.5	11.4	5889	2	US-08-462-109A-5	Sequence 5, Appli
40	687.5	11.4	5889	3	US-08-460-907B-5	Sequence 5, Appli
41	687.5	11.4	5889	3	US-08-463-179A-5	Sequence 5, Appli
42	687.5	11.4	5889	3	US-08-461-384B-5	Sequence 5, Appli
43	674.5	11.2	5120	3	US-08-772-270A-6	Sequence 6, Appli
44	674.5	11.2	8370	2	US-08-488-706-1	Sequence 1, Appli
45	671.5	11.2	5011	1	US-08-463-092B-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1

US-08-181-471-2
; Sequence 2, Application US/08181471
; Patent No. 5641508
; GENERAL INFORMATION:
; APPLICANT: Li, Lingna
; APPLICANT: Lishko, Valeryi K.
; TITLE OF INVENTION: METHOD FOR DELIVERING BENEFICIAL
; TITLE OF INVENTION: COMPOSITIONS TO HAIR FOLLICLES
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Thomas Fitting
; STREET: 12526 High Bluff Drive, Suite 300
; CITY: San Diego
; STATE: CA
; COUNTRY: USA
; ZIP: 92130
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/181.471
; FILING DATE: 13-JAN-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/041,553
; FILING DATE: 02-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: ANT0029P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-792-3680
; TELEFAX: 619-792-8477

; INFORMATION FOR SEQ ID NO: 2:

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4646 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 425..4267
; US-08-181-471-2
```

Alignment Scores:

```
Pred. No.: 0 Length: 4646
Score: 3545.00 Matches: 681
Percent Similarity: 75.55% Conservative: 246
Best Local Similarity: 55.50% Mismatches: 264
Query Match: 58.97% Indels: 36
DB: 1 Gaps: 9
```

US-09-873-409-6 (1-1195) x US-08-181-471-2 (1-4646)

```
QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
DB 575 ATGGTGGTGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCTCTCATGATCGTG 634
QY 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
DB 635 GTGTTTGGAGAAATGACAGATATCTTTGCAAAATGCGAGAAATTTAGAAATGATCGATGTC 694
QY 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
DB 695 AACATCACTAATAGAGTATGATCAATGATACAGGGTCTTCATGATCTGGAGGAGAC 754
QY 44 -----ArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
DB 755 ATGACCAAGGTATGCCATTATTACAGTGGAAATGGTCTGGGGTGGTGGTGGCTTAC 814
QY 62 IleGlnIleSerLeuThrIleThrAlaAlaArgGlnThrIysArgIleArgLysGln 81
DB 815 ATTCAGGTTTCATTTGGTGGCTGGCAGCTGGGAAGACAAATACACAAAATTAGAAACAG 874
QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
DB 875 TTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTGATGTCACGATGTTGGGGAG 934
QY 102 LeuAsnThrArgMetThr---AspIleAspIleSerAspGlyIleGlyAspLysIle 120
DB 935 CTTAACAACCCGAGCTTACAGATGATGCTCTTAAGATTAAATGAGTTATTGGTACAAAAT 994
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
DB 995 GGAATGTTCTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGGATTACACGT 1054
QY 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
DB 1055 GGTGGGAAGCTTAACCTGTGATTTTGGCCATCATGCTCTGTCTTGGACATGTCAGCTGCT 1114
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
DB 1115 GTCTGGCCAAAGATACCTATCTTCAATTTACTCATTAAGAACTCTTAGCGTATGCCAAGCT 1174
QY 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
DB 1175 GGAGCAGTAGCTGAAGAGGCTCTGGCAGCAATTAGAACTGTGATTCATTTGGAGGACAA 1234
QY 201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
DB 1235 AAGAAAGAACTTGAAGGTACAAACAAATTTAGAAAGCTAAAGAAATTTGGGATTAAG 1294
QY 221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
```

```
DB 1295 AAAGCTATTACAGCCAAATATTCTATAGTGTGCTTCTCTGCTGATCATCTTAT 1354
QY 241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
DB 1355 GCTCTGGCCCTTCTGTGTATGGGACCACTTGGTCCCTCTCAGGGGAA-----TATTCTATT 1408
QY 261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
DB 1409 GGACAGTACTCACTGTATCTTCTTCTGTATTAAATGGGGCTTTAGTGTGGCAGGCA 1468
QY 281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
DB 1469 TCTCCAAGCATTTGAAGCATTTGCAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1528
QY 301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
DB 1529 ATTGATAATAAGCAAGTATTGACAGCTATTGAGAGGTGGGACAAACCAAGATAATT 1588
QY 321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
DB 1589 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAGTTAAG 1648
QY 341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
DB 1649 ATCTTGAAGGGCTGAACCTGAAGGTGACAGTGGGACAGCGTGGCCCTGGTGGAAAC 1708
QY 361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
DB 1709 AGTGGGTGGGGAAGAGCACACAGCTCCAGCTGATGACAGGCTCTATGACCCACAGAG 1768
QY 381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
DB 1769 GGGATGGTCAGTGTGATGGACAGGATATTAGACCATAAATGTAAGGTTTCTACGGGAA 1828
QY 401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
DB 1829 ATCATTGGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1888
QY 421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
DB 1889 CGCTATGGCCGTAAGAAATGTCACCATGGATGAGATTTGAGAAAGCTGTCAAGGAGGCAAT 1948
QY 441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
DB 1949 GCCTATGACTTTATCATGAACTGCCTCATAAATTTGACACCTGCTGGTGGAGAGAGGG 2008
QY 461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
DB 2009 GCCCAGTTGAGTGGTGGGAGAGCAGAGAGATCGCCATTTGCAGCTGCCCTGGTTCGCAAC 2068
QY 481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
DB 2069 CCCAAGATCTCTGCTGGATGAGCCAGCTCAGCTTGGACACAGAAAGCGAGAGTGT 2128
QY 501 ValGlnAlaAlaLeuGluLysAlaSerLysGlyValGlyThrThrIleValValAlaHisArg 520
DB 2129 GTTCAGTGGCTCTGATTAAGGCCAGAAAGGTGGGACCACTTGTGATGATGCTCATCGT 2188
QY 521 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 540
DB 2189 TTGCTACAGTTCGTAAATGCTGACGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2248
QY 541 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 560
DB 2249 AAGGAAATCATGATGAATCATGAAAGAAAGGAGGATTTTACTTCAAACTTGTGCAAAATG 2308
QY 561 Gln-----AspIleLysLysAlaAspGluGlnMetGluSerMetThr 574
DB 2309 CACAGCAGCAGAAATGAAGTTGAATTAGAAATGACAGCTGATGATCAATCCAAAGTGAAT 2368
QY 575 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal----- 589
```


/ COMPUTER: IBM PC compatible
 / OPERATING SYSTEM: PC-DOS/MS-DOS
 / SOFTWARE: Patent In Release #1.0, Version #1.30
 / CURRENT APPLICATION DATA:
 / APPLICATION NUMBER: US/08/793,610
 / FILING DATE: 07-MAR-1997
 / PRIOR APPLICATION DATA:
 / APPLICATION NUMBER: DE P 44 31 973.8
 / FILING DATE: 08-SEP-1994
 / PRIOR APPLICATION DATA:
 / APPLICATION NUMBER: DE 195 03 952.1
 / FILING DATE: 07-FEB-1995
 / APPLICATION DATA:
 / APPLICATION NUMBER: PCT/BP95/03175
 / FILING DATE: 10-AUG-1995
 / ATTORNEY/AGENT INFORMATION:
 / NAME: Berman, Richard J.
 / REGISTRATION NUMBER: 39,105
 / REFERENCE/DOCKET NUMBER: P1614-7007
 / TELECOMMUNICATION INFORMATION:
 / TELEPHONE: (202) 638-5000
 / TELEFAX: (202) 638-4810
 / INFORMATION FOR SEQ ID NO: 5:
 / SEQUENCE CHARACTERISTICS:
 / LENGTH: 6505 base pairs
 / TYPE: nucleic acid
 / STRANDEDNESS: double
 / TOPOLOGY: circular
 / MOLECULE TYPE: DNA
 / US-08-793-610-5

Alignment Scores:

Pred. No.: 0 Length: 6505
 Score: 3545.00 Matches: 681
 Percent Similarity: 75.55% Conservative: 246
 Best Local Similarity: 55.50% Mismatches: 264
 Query Match: 58.97% Indels: 36
 DB: 2 Gaps: 9

US-09-873-409-6 (1-1195) x US-08-793-610-5 (1-6505)

QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
 Db 1967 ATGGTGGTGGAACTTTGGCTGGCCATCATCCATGGGCTGGACTTCTCTCATGATGCTG 2086
 QY 21 ValIleuGlyGluMetSerAsp-----AsnLeuIleSer 31
 Db 2027 GTGTTGGAGAAATGACAGATATCTTTGCAAAATGCGAGAAATTTAGAGATCTGATGTCA 2086
 QY 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
 Db 2087 AACATCACTAATAGAGTGTATCATGATACAGGGTTCTTCATGAATCTGGAGGAGAC 2146
 QY 44 -----ArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
 Db 2147 ATGACCAAGGTATGCTATTATTACAGTGAATGGTCTGGGGTGGTGGTGGCTTAC 2206
 QY 62 IleGlnIleSerLeuThrIleThrAlaAlaArgGlnThrIleArgIleArgGlyGln 81
 Db 2207 ATTCAAGTTTCATTTTGGTGGCTGGAGCTGGAGACAAATACACAAATTTAGAAACAG 2266
 QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101
 Db 2267 TTTTTCATGCTATATGACAGAGGATAGGCTGTTGATGTGCACGATGTGGGAG 2326
 QY 102 LeuAsnThrArgMetThr-----AspIleAspIleSerAspGlyIleGlyAspIle 120
 Db 2327 CTTAAACACCCGACTTACAGATGATGCTCTTAAGATTAATCAAGTTATTGTCGACAAAT 2386
 QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
 Db 2387 GGAATGTTCTTCAGTCAATGGCAACATTTTTCAGTGGGTTTATAGTAGGATTTACACGT 2446

QY 141 GlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
 Db 2447 CGTTGGAGAGCTAACCTTGTGATTTGGCCATCAGTCTCTGTTCTTGGAGTGTGAGTGT 2506
 QY 161 AlaCysSerArgMetValIleSerLeuThrSerIleGluLeuSerAlaTyrSerLysAla 180
 Db 2507 GTCTGGGCAAGATACTATCTTCATTCTGATAAGAACTCTTAGCGTATCAAAAGCT 2566
 QY 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
 Db 2567 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGACTGTGATTCATTCGATTGGAGCAA 2626
 QY 201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyLys 220
 Db 2627 AAGAAAGAACTTGAAGAGGTACAAACAAATTTAGAAGAACTTAAAGAAATTTGGATAAG 2686
 QY 221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
 Db 2687 AAAGCTATTACAGCCAATATTCTATAGTGTGCTTCTCTGCTGCTATCATCTCTTAT 2746
 QY 241 GlyLeuAlaPheTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
 Db 2747 GCTCTGGCCCTTCTGTTATGGGACCACTTGGCTCTCAGGGGAA-----TATTCAT 2800
 QY 261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
 Db 2801 GCACAAGTACTCACTGATTTCTTTCTGTATTAAATTTGGGGCTTTTAGTGTGGACAGGCA 2860
 QY 281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
 Db 2861 TCTCCAAGCATTTGAAGCATTTTGCATAATGCAAGAGGAGCACTTATGAAATCTTCAAGATA 2920
 QY 301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
 Db 2921 ATGTATATTAACCAAGTATTGACAGCTATTTCAGAGTGGGACACAAACCAATATAT 2980
 QY 321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerLys 340
 Db 2981 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCACTTCACTTCACTTCACTTCACTTCA 3040
 QY 341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
 Db 3041 ATCTTGAAGGGCTGAACCTGAAGGTGACAGTGGGACAGCGTGGCCCTGTTGTAAC 3100
 QY 361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAsp 380
 Db 3101 AGTGGCTGTGGAGAGGACACACAGTCCAGCTGATGCAGAGCTCTATGACCCACAGAG 3160
 QY 381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
 Db 3161 GGGATGCTCAGTGTGATGGACAGGATATTAGGACCATAAATGTAAGTTCCTACGGAA 3220
 QY 401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
 Db 3221 ATCAATTTGGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 3280
 QY 421 LysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
 Db 3281 CGCTATGGCCGTGAAATGTACCATGATGATGATGATGATGATGATGATGATGATGATGAT 3340
 QY 441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGly 460
 Db 3341 GCCTATGACTTTTATCATGAACTGCTCTATAAATTTGACACCTTGTGAGAGAGAGGG 3400
 QY 461 AlaGlnMetSerGlyGlyGlnIleGlnAlaIleAlaArgAlaLeuValArgAsn 480
 Db 3401 GCCCAGTTTCAAGTGGTGGGAGGAGGAGGATCGCCATTCGACGTGCCCTGTTGTCACAC 3460
 QY 481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
 Db 3461 CCCAAGATCTCTGCTGATGAGCCACGTGACCTTGGACACAGAAAGCCAGAGGTG 3520
 QY 501 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 520

QY 421 LysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
DB 3240 CGTATGGCGTGAAGATGTCACCATGGATGAGATTGAGAAAGCTTCAGAAAGCAAT 3299
QY 441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
DB 3300 GCTATGACTTATCATGAACTGCTCATTAATTTGACCCCTGGTTGGAGAGAGGG 3359
QY 461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
DB 3360 GCCCAGTTGAGTGGCGAGAGAGAGGATCGCATTCGACGTGCCCTGGTTGCCAAC 3419
QY 481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaIleAspSerGluSerLysSerAla 500
DB 3420 CCCAAGATCCTCTGCTGGATGAGGCACGTCAGCCTTGGACACAGAAAGCGAAGCATG 3479
QY 501 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 520
DB 3480 GTTCAGTGGCTCTGGATAGGCCAGAAAGGTCGACCAACCATTTGTATGATCTCATCGT 3539
QY 521 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 540
DB 3540 TTGCTCTACAGTTGCTAATGCTGACGTCATCGCTGTTTCGATGATGGAGTCATTGGGAG 3599
QY 541 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 560
DB 3600 AAAGGAATCATGATGAACCTCATGAAAGGAGAAAGGCAATTTACTTCAAACTTGTCACAATG 3659
QY 561 Gln-----AspIleLysLysAlaAspGluGlnMetGluSerMetThr 574
DB 3660 CAGACAGCAGGAATGAAGTTGAATTAGAAATGCGCATGATGAATCCAAAGAGTGAAT 3719
QY 575 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal----- 589
DB 3720 GATGCTTGGAAATGCTTCAATGATTCAGATCCAGTCTAATAGAAAAGATCAACT 3779
QY 590 ---LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys 607
DB 3780 CGTAGGATGTCGCTGGATCACAAGGCCAAGACAGAAAGCTTAGTACCAAGAGGCTCTG 3839
QY 608 GluIleSerLeuProGluValSerLeuLysIleLysLysLeuLysProGluTyr 627
DB 3840 GATGAAGATATACCTCCAGTTTCTTTGGAGGATTATGAAGCTAAATTTAACTGAATGG 3899
QY 628 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 647
DB 3900 CCTATTATTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 3959
QY 648 SerIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 666
DB 3960 GCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGAAATTCATGATCCTGAAACAAA 4019
QY 667 LysHisAspAlaGluIleLysSerMetIlePheValIleLeuGlyValIleCysPheVal 686
DB 4020 CGACAGAAATAGTAACTGTTTTCACATTTGTTCTAGCCCTTGAATATTTCTTTTATT 4079
QY 687 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyLysIleLeuThrMetArgLeu 706
DB 4080 ACATTTTCTCTCAGGGTTTCATTTGGCAAGAGCTGGAGAGATCTCCACCAAGCGGCTC 4139
QY 707 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaIlePheAspGluLysGlu 726
DB 4140 CGATACATGTTTTCGATCCATGCTCAGACAGGATGTGAGTTGGTTGATGACCTATAA 4199
QY 727 AsnSerThrGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 746
DB 4200 AACACCACTGGACATGACTACAGCTCGCAATGATGCTCTCAAGTTAAAGGGGCT 4259
QY 747 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 766
DB 4260 ATAGGTTCCAGGCTGCTGTGAATTTCCAGAAATATAGCAAAATCTTGGGACAGGAAT 4319
QY 767 IleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeu 786

DB 4320 ATATCTCTTATCTATGTTGGCAACTTAACACTGTTACTCTTAGCAATTTACCCATCAT 4379
QY 787 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 806
DB 4380 GCATAGCAGGAGTGTGTAATGAAATGTTGTTCTGGACAGCACTGAAGATAGAAA 4439
QY 807 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 826
DB 4440 GAACTAGAAGGTCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCGCAACCGTTGT 4499
QY 827 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 846
DB 4500 TCTTTGACTCAGAGCAGAAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATAC 4559
QY 847 ArgAsnThrSerLysLysAlaGlnIleGlySerCysTyrAlaPheSerHisAlaPhe 866
DB 4560 AGAACTCTTTGAGGAAAGCACACATCTTTGGAAATACATTTCTTCCACCCAGGCAATG 4619
QY 867 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 886
DB 4620 ATGTATTTTCTATGCTGATGTTTCCGTTTGGAGCTTACTTGTGTGCACATAAATC 4679
QY 887 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 906
DB 4680 ATGAGCTTTGAGGATGTTCTGTAGTATTTTTCAGCTGTTGTTGTTGTTGTTGTTGTTG 4739
QY 907 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaHisLeu 926
DB 4740 GGGCAGTCACTGTTCTGCTGATGATGCCAAAGCCAAATATATCAGACGCCACATC 4799
QY 927 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluLysLysPro 946
DB 4800 ATCATGATCATTTGAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGSCCTAATGCCG 4859
QY 947 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 966
DB 4860 AACACATTTGGAAGAAATGTCATTTGTTGTAAGTTGTTATTCAACTATATCCACCCGACCG 4919
QY 967 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 986
DB 4920 GATATCCAGTCTCTCAGGACTGAGCTGGAGGTGAAAGGCCAGACGCTGGCTCTG 4979
QY 987 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuGlnArgLeuTyrAsp 1006
DB 4980 GTGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCTCTGGAGCGGTTCTACGAC 5039
QY 1007 ProValGlnGlyClnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTyr 1026
DB 5040 CCTTTGGCAGGGAAGTGTCTGTGTCAAAGAAATAAAGCGACTGAATGTTCAAGTGG 5099
QY 1027 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 1046
DB 5100 CTCCGAGCACACTGGGCATCTGTCCTCCAGGAGCCCATCTCTGTTTACCTGACATGCT 5159
QY 1047 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 1066
DB 5160 GAGAACATTTGCTATGGACACAAACAGCCGGTGTGTCTACAGGAAGAGATCGTGAGGCA 5219
QY 1067 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 1086
DB 5220 GCAAGAGGAGGCCAACATACATGCTTTCATCGAGTCACTGCCTTAATAAATATAGCACTAA 5279
QY 1087 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlnLysGlnArgLeuAlaIleAlaArg 1106
DB 5280 GTAGGAGACAAAGGAACCTCAGCTCTCTGTGTGGCCAGAAACAACGATATGCCATGCTCGT 5339
QY 1107 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 1126
DB 5340 GCCTTTGTAGACGCTCATATTTTGTGTTTGGATGAAGCCACGTCAGCTCTCGATACA 5399
QY 1127 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1146
:::|||||

Db 5400 GAAAGTGAAGAGGTGTCTCAAGAGCCCTGGACAAAGCCGACAGAGCCGACCTGCATT 5459
 QY 1147 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1166
 Db 5460 GTGATGTCTCACCCTGTCTCCACATCCAGATTCAGACTTAATAGTGGTGTTCAGAAAT 5519
 QY 1167 GlyLysIleLysGluGlnGlyThrHisGlnGlnLeuLeuArgAsnArgAspIleTyrPhe 1186
 Db 5520 GGCAGAGTCAAGGAGCATGCGCAGCATCAGCAGTGTCTGGCACAGAAAGGACCTATT 5579
 QY 1187 LysLeuValAsnAlaGlnSer 1193
 Db 5580 TCAATGGTCAAGTGTCCAGGCT 5600
 RESULT 4
 US-08-784-649A-1
 ; Sequence 1, Application US/08784649A
 ; Patent No. 5830697
 ; GENERAL INFORMATION:
 ; APPLICANT: Sikic, Branimir I
 ; APPLICANT: Chen, Gang
 ; TITLE OF INVENTION: P-GLYCOPROTEIN MUTANT RESISTANT TO
 ; CYCLOSPORIN MODULATION
 ; NUMBER OF SEQUENCES: 5
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Fish & Richardson
 ; STREET: 2200 Sand Hill Road
 ; CITY: Menlo Park
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94025
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/784,649A
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Sherwood, Pamela J
 ; REGISTRATION NUMBER: Reg.No. 5830697 36,677
 ; REFERENCE/DOCKET NUMBER: 06037/007001
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415-322-5070
 ; TELEFAX: 415-854-0875
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 4264 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 US-08-784-649A-1
 Alignment Scores:
 Pred. No.: 0 Length: 4264
 Score: 3543.50 Matches: 682
 Percent Similarity: 75.63% Conservative: 246
 Best Local Similarity: 55.58% Mismatches: 262
 Query Match: 58.94% Indels: 37
 DB: 2 Gaps: 10
 US-09-873-409-6 (1-1195) x US-08-784-649A-1 (1-4264)
 QY 1 MetIleuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
 Db 292 ATGGTGGGGAACCTTTGGTGGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 351
 QY 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
 Db 352 GTGTTTGGAGAAATGACAGATATCTTTGCAAAATGACAGAAATTTAGAAAGATCTGATGTCA 411

QY 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
 Db 412 AACATCACTAATAAGAGTGATATCAATGATACAGGTTCTTTCATGAATCTGGAGGAAC 471
 QY 44 -----ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
 Db 472 ATGACACAGGTATCCCTATTATTACAGTGGAAATGGTGTGGGGTGTCTGTGTGCTTAC 531
 QY 62 IleGlnIleSerLeuTyrIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
 Db 532 ATTCAGTTTTCATTTTGGGCTCGCAGCTGGAGACAAATACACAAAATTTAGAAACAG 591
 QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101
 Db 592 TTTTTCATGCTAATAATGCGACAGGAGATAGGTGGTTGATGTGCACGATGTGGGAG 651
 QY 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
 Db 652 CTTAACACCCGACTTACAGATGATGTCTCAAGATTAATGAAGGAATTTGGTGACAAAAT 711
 QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
 Db 712 GGA---ATGTTTCAGTCAATGGCAACATTTTCACCTGGGTTTATAGTAGGATTTACCGT 768
 QY 141 GlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
 Db 769 GGTGGAGACTAACCTTGTGATTTGGCCATCAGTCTCTGTTCTGGACTGTGAGTGTGT 828
 QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
 Db 829 GTCTGGGCAAGATACATCTCTTCTTACTATAAAGAACTCTTAGCGTATGCAAAAGCT 888
 QY 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
 Db 889 GGAGCAGTAGCTGAAGAGCTCTGGCAGCAATTAGCAACTGTGATTCATTTGGAGGACAA 948
 QY 201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
 Db 949 AGAAGAAGACTTGAAGGTACAAACAAATTTAGAAGAGCTAAGAAGATTTGGATAG 1008
 QY 221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
 Db 1009 AAAAGCTATTACAGCAATATTCTATAGTGTCTGCTCTCTGCTGATCTATCATCTTAT 1068
 QY 241 GlyLeuAlaPheTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
 Db 1069 GCTCTGGCCTTCTGGTATGGGACCCTTGGTCTCTCAGGGGAA-----TATTCTATT 1122
 QY 261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
 Db 1123 GGACAGTACTCCTGTTATTTCTGTATTAAATGGGGCTTTTAGTGTGGACAGGCA 1182
 QY 281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
 Db 1183 TCTCCAAGCATTGAAGCATTTCAAATGCAAGAGGAGCAGCTTATCAAAATCTTCAAGATA 1242
 QY 301 IleAspLysIleProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
 Db 1243 ATTGATAATAAGCAAGTATTGACAGCTATTGCAAGAGTGGGCGACAAACAGATAATATT 1302
 QY 321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
 Db 1303 AAGGGAATTTGGAAATTCAGAAATGTTCTCAGTTACCATCTCCAAAGAGTTAAG 1362
 QY 341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
 Db 1363 ATCTTGAAGGCTGAACCTGAAGGTGCAGAGTGGCGACAGCGTGGCCCTGGTTGAAAC 1422
 QY 361 AsnGlySerGlyLysSerThrValValGlnLeuGlnArgLeuTyrAspProAspAsp 380
 Db 1423 AGTGGTGTGGGAAGAGCAACAGTCCAGCTGATGCAGAGGCTCTATGACCCACAGAG 1482

QY 381 GlyPheIleMetValAspGluLeuAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
DB 1493 GGGATGTCAGTGTGATGGACAGAGATATATAGGACCAATAATGTATAGGTTTCTACGGAA 1542
QY 401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
DB 1543 ATCATTTGGTGGTGGAGTCAGGAACCTGTATTTGTCACACAGATAGCTGAACAAT 1602
QY 421 LysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
DB 1603 CGCTATGGCCGTGAATAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCAAT 1662
QY 441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyLysGly 460
DB 1663 GGCATGACATTTATCATGAACTGCCTCAATAATTTTGACACCCCTGGTTGGAGAGAGGG 1722
QY 461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
DB 1723 GCCCAGTTGAGTGGGCGAAGACAGAGATCGCCATTTGCAGCTGCTGTTTGGCAAC 1782
QY 481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
DB 1783 CCCAAGATCCTCTGCTGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGT 1842
QY 501 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 520
DB 1843 GTTCAGGTGGCTGTGATAGGCCAGAAAGGTGCGACCACCATTTGTGATAGCTCATCGT 1902
QY 521 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 540
DB 1903 TTGCTACAGTTGCTGTAATGCTGACGTCATCGCTGTTTCGATGATGGATGCATTTGGAG 1962
QY 541 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 560
DB 1963 AAAGGAATCATGATGAATCATGAAAGAGAAAGGCAATTTACTTCAAACTGTGCACAATG 2022
QY 561 Gln-----AspIleLysLysAlaAspGluGlnMetGluSerMetThr 574
DB 2023 CACACAGCAGGAAATGAAGTTGAATTAGAAATGCGAGCTGATGAATCCAAAGTGAAAT 2082
QY 575 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal----- 589
DB 2083 GATGCCCTGGAAATGCTTCAATAGATTCAAGATCCAGTCTAATAGAAAAAGATCAACT 2142
QY 590 ---LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys 607
DB 2143 CGTAGAGTGTCCGTGGATCAGACAGCCCAAGACAGAAAGCTTAGTACCAAGAGGCTCTG 2202
QY 608 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTyr 627
DB 2203 GATGAAAGTATACCTCCAGTTTCCCTTTTGGAGGATATGAAGCTAAATTTAACTGAATGG 2262
QY 628 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 647
DB 2263 CCTATTTTGTGGTGATTTTGTGTCATTTATAATGGAGGCTGCAACAGCATTT 2322
QY 648 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 666
DB 2323 GCAATAATATTTTCAAGATATATAGGGTTTTACAGAAATGTATGATCTCTGNAACAAA 2382
QY 667 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 686
DB 2383 CGACAGATAGTAACATGTTTTCACATTTTGTAGCCCTTGGAAATTTATTTCTTTATT 2442
QY 687 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 706
DB 2443 ACATTTTCTTCAGGGTTTCAATTTGGCAAGCTTGGAGAGATCTCTCAACAGCGGCTC 2502
QY 707 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGluLysGlu 726
DB 2503 CGATACATGGTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTGATGACCCCTAAA 2562
QY 727 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 746

DB 2563 AACACACTGGAGCATTTGACTACCAAGGCTCGCAATGATGCTGCTCAAGTTAAAGGGGT 2622
QY 747 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 766
DB 2623 ATAGGTTCCAGGCTTGCTGTAATTTACCCAGAAATATAGCAAAATCTTGGGACAGAAAT 2682
QY 767 IleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeu 786
DB 2683 ATATCCTTTCATCTATGTTGGCAACTAAACACTGTTTACTCTTAGCAATTTGACCCATCA 2742
QY 787 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 806
DB 2743 GCAATAGCAGGAGTTGTTGAAATGAAATGTTGCTGGCAAGACACTGAAGATATAAGAA 2802
QY 807 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 826
DB 2803 GAACTAGAAGGTCCTGGAGATCGCTACTGAAGCAATAGAAAACCTTCGACACCGTTGT 2862
QY 827 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 846
DB 2863 TCTTTGACTCAGGACAGAAAGTTTGAACATATGATGCTCAGAGTTTGCAAGTACCATAC 2922
QY 847 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 866
DB 2923 AGAAACTCTTTGAGGAAACACACATCTTTGGAATTTACATTTTCTTCCACCGGCAATG 2982
QY 867 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyValAlaTyrLeuIleGlnAlaGlyArg 886
DB 2983 ATGATTTTTCATGCTGGATGTTCCGTTTGGAGCTACTTGTGGCACATAAATC 3042
QY 887 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 906
DB 3043 ATGAGCTTTGAGGATGTTCTGTTAGTTATTTTTCAGCTGTTGTTGTTGTTGTTGTTG 3102
QY 907 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaHisLeu 926
DB 3103 GGGCAAGTCAGTTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3162
QY 927 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysPro 946
DB 3163 ATCATGATCATTTGAAATAAACCCCTTTGATTGACAGTACAGCACGGAGGCTTAATGCCG 3222
QY 947 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 966
DB 3223 AACACATTGGAGGAAATGTCACATTTGCTGAAGTTGTTATTCAACTATCCACCCGACCG 3282
QY 967 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 986
DB 3283 GACATCCCAAGTCTTTCAGGACTGAGCTGAGGCTGAGAGGCGGACGCTGGCTCTG 3342
QY 987 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuGlnArgLeuTyrAsp 1006
DB 3343 GTGGCAGCAGTGGCTGTGGGAAGACACAGTGGTCCAGCTCTGAGCGGTTCTACGAC 3402
QY 1007 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTyr 1026
DB 3403 CCTTTGGCAGGAAAGTGTCTGCTGATGTCGCAAGAAATAAAGCGACTGAATGTTCACTG 3462
QY 1027 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 1046
DB 3463 CTCGAGCAGCACCTGGGCATCGTCCAGGAGGCCATCTCTGTTGACTGACGATTCGT 3522
QY 1047 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 1066
DB 3523 GAGAACATTCCTATGGAGACAAACAGCCGGGGTGTCCACGGAAGAGATTTGTGGGGCA 3582
QY 1067 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 1086
DB 3583 GCAAGGAGGCCAAATACATGCTTCATCGAGTCACTGCTTAATAATATATAGCACTAAA 3642
QY 1087 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 1106

Db 3643 GTAGGAGACAAAGAACTCAGCTCTCTGTGGCCAGAAAACAGCATTGCCATAGCTCGT 3702
QY 1107 AlaleuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 1126
Db 3703 GCCTTGTAGACAGCTCATATTTTGTCTTTGGTGAAGCCAGCTCAGCTCTGGATACA 3762
QY 1127 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1146
Db 3763 GAAAGTGAAAGGTTGTCCAAGAGCCCTGGACAAAGCCAGAGAGCCGACCTGCATT 3822
QY 1147 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1166
Db 3823 GTATGCTCACCGCTGTCCACATCCAGATCCAGACTTAATAGTGGTGTTCAGAAAT 3882
QY 1167 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1186
Db 3883 GGCAGAGTCAAGAGCATGCGACGATCAGCAGCTGCTGCGACAGAAAGCAATCTATTTT 3942
QY 1187 LysLeuValAsnAlaGlnSer 1193
Db 3943 TCAATGCTCAGTGTCCAGGCT 3963

RESULT 5

US-08-784-649A-5
; Sequence 5, Application US/08784649A
; Patent No. 5830697
; GENERAL INFORMATION:
; APPLICANT: Sikic, Branimir I
; APPLICANT: Chen, Gang
; TITLE OF INVENTION: P-GLYCOPROTEIN MUTANT RESISTANT TO
; TITLE OF INVENTION: CYCLOSPORIN MODULATION
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 2200 Sand Hill Road
; CITY: Menlo Park
; STATE: CA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/784,649A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: Reg.No. 5830697 36,677
; REFERENCE/DOCKET NUMBER: 06037/007001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-322-5070
; TELEFAX: 415-854-0875
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4264 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-784-649A-5
Alignment Scores:
Pred. No.: 0 Length: 4264
Score: 3543.50 Matches: 682
Percent Similarity: 75.63% Conservative: 246
Best Local Similarity: 55.58% Mismatches: 262
Query Match: 58.94% Indels: 37
DB: 2 Gaps: 10

US-09-873-409-6 (1-1195) x US-08-784-649A-5 (1-4264)

QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 292 ATGGTGGTGGGAACCTTGGCTGCCATCATCCATGGGGCTGGAGCTTCCTCTCATGATGCTG 351
QY 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
Db 352 GTGTTTGGGAAATGACAGATATCTTGGCAATGTCAGGAAATTTAGAAGATCTGATGTCA 411
QY 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
Db 412 ACATCAGCTAATAGAGTGCATATCAATGATACAGGTTCTTTCATGAATCTGGAGGAAGAC 471
QY 44 -----ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db 472 ATGACCCAGGTATGCCCTATTATTACAGTGAATTTGGTGGGTGCTGGTGTGCTGTAC 531
QY 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db 532 ATTCAAGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAATTTAGAAACAG 591
QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101
Db 592 TTTTTCATGCTATATATGACAGGAGATAGGCTGTTTGTATGTGCACCATGTTGGGAG 651
QY 102 LeuAsnThrArgMetThr-----AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db 652 CTTAACACCCCGACTTACAGATGATGTCCTCAAGATTAATGAAGGAATTTGGTGACAAAT 711
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db 712 GGA-----ATGTTCCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGGATTACAC 768
QY 141 GlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db 769 GGTGGAAGCTAACCTTGTGATTTGGCCATCAGTCTCTTCTTGGACTGTGACGTGCT 828
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db 829 GTCTGGCAAGATATACTATCTTCTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 888
QY 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db 889 GGAGCAGTAGCTCAAGAGGCTCTTGGCAGCAATTAGAACCTGTGATTCATTTGGAGACAA 948
QY 201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
Db 949 AGAAAGAACTTGAAAGGTACACAAATAATTTAGAAGAGCTAAAGAAATTTGGGATAAG 1008
QY 221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
Db 1009 AAAGCTATTACAGCCAATATTCTATAGTGTGCTTCTTCTGCTGATCTATGATCTTAT 1068
QY 241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
Db 1069 GCTCTGGCCTTCTGGTATGGGACCACTTGGTCTCTCAGGGGAA-----TATTCTATT 1122
QY 261 GlyThrValIleAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
Db 1123 GGACAAGTACTCACTGTATTCTTCTGTATTAAATTTGGGGCTTTTAGTGTGTGACAGGCA 1182
QY 281 ValProHisPheGluThrPheAlaIleAlaArgGlyValAlaPheHisIlePheGlnVal 300
Db 1183 TCTCCAAGCATTTGAACATTTGCAATGCAAGAGGAGCAGCTTATGAATCTTCAAGATA 1242
QY 301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
Db 1243 ATTGATAATAAGCCAAAGTATTGACAGCTATTGACAGAGTGGGACAAACACAGATAATAT 1302
QY 321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
Db 1303 AAGGGAAATTTGGAAATTCAGAAATGTTTCACTTCACTTCCATCTCGAAAGAGTTAAG 1362

Qy	341	IleuLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu	360
Db	1363	ATCTTGAAGGCCCTGAACTCGAAGTGCAGAGTGGCAGACGGTGGCCCTGGTTGGAAAC	1422
Qy	361	AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyArgProAspAsp	380
Db	1423	AGTGGCTGTGGGAAGACCAACAGTCCAGCTGATGAGAGGCTCTATGACCCACAGAG	1482
Qy	381	GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyArgAsp	400
Db	1483	GGGATGGTCAGTGTTCATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA	1542
Qy	401	HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle	420
Db	1543	ATCATTTGGTGTGAGTCAGGAACCTGTATTGTTGTCGCCACCATAGCTAGTAAGAAACATT	1602
Qy	421	LysTyArgArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn	440
Db	1603	CGCTATGGCCGTGAATAATGTCACCATGGATGAGATTGAGAAGCTGTCAAGGAAGCCCAAT	1662
Qy	441	AlaTyArgPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly	460
Db	1663	GCCTATGACCTTATCATGAAACCTGCCTCATAAATTGACACCCCTGGTGAGAGAGAGGG	1722
Qy	461	AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn	480
Db	1723	GCCCACTTGAGTGGTGGCAGAACGACAGATCGCCATTGCACGTGCCCTGGTTGCGCAAC	1782
Qy	481	ProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla	500
Db	1783	CCCAAGATCTCTCTGTGGATGAGGCCACGTGAGCTTGGACACAGAAGCGAAGCAGTGT	1842
Qy	501	ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg	520
Db	1843	GTTCCAGTGGCTCTGGATAGGCCAGAAAGGTGGGACCAACCATTTGCTATAGCTCATCGT	1902
Qy	521	LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu	540
Db	1903	TTGTCTACAGTTCGTATGCTGACGTCATCGCTGGTGTTCGATGATGGAGTCATTTGGGAG	1962
Qy	541	LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyTyThrSerLeuValMetSer	560
Db	1963	AAAGGAATCATGATGAACTCATGAAGAAGAAGGCATTTACTTCAAACTTGTCCAAATG	2022
Qy	561	Gln-----AspIleLysLysAlaAspGluGlnMetGluSerMetThr	574
Db	2023	CAGACAGCAGGAATGAAGTTGAATTAGAAATGACGCTGATGATCAAAAGTGAAATT	2082
Qy	575	TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal-----	589
Db	2083	GATCCCTTGGAAATGCTCTCAAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACT	2142
Qy	590	---LysSerLysLys---SerAspPheIleAspLysAlaGluLysSerThrGlnSerLys	607
Db	2143	CGTAGGAGTGTCCGTGTGATCACAAGCCCAAGACAGAAAGCTTTAGTACCAAGAGCGCTGTG	2202
Qy	608	GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTyr	627
Db	2203	GATGAAGTAGTACCTCCAGTTTCTTTTGGAGGATTTATGAGCTAAATTTAACTGAATGG	2262
Qy	628	ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe	647
Db	2263	CCTTATTTTGTGTGTGTTATTTTGTGCCATTATATAATGGAGGCCCTGCAACCAGCATTT	2322
Qy	648	SerIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu	666
Db	2323	GCAATAATATTTTCAAAAGATATATGAGGTTTTCACAAAGAAATTCATGATCCTGCAACAAA	2382
Qy	667	LysHisAspAlaGluIleTySerMetIlePheValIleLeuGlyValIleCysPheVal	686
Db	2383	CGACAGAAATAGTAACCTGTTTTCACATATGTTTCTAGACCTTGGAAATTTATTTCTTTTAT	2442
Qy	687	SerTyPheMetGlnGlyLeuPheTyArgAlaGlyGluIleLeuThrMetArgLeu	706

Db 3523 GAGAACATTGCCTATGAGAGCAACAGCCGGTGGTGTGCACAGGAAGAGATTGTGAGGCA 3582
Qy 1067 AlaAsnAlaalaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 1086
Db 3583 GCAAGAGGCGCAACATACATGCTTTCATCGAGTCACTGCCTTAATAATATAGCACTAAA 3642
Qy 1087 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 1106
Db 3643 GTAGGACAAAGAACTCAGCTCTCTGGTGGCCAGAAACAGCATGCGCTAGCTCGT 3702
Qy 1107 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 1126
Db 3703 GCCCTTTAGACAGCCTCATATTTTCTTTTGGATGAAGCCACGTCAGCTCTGGATACA 3762
Qy 1127 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1146
Db 3763 GAAAGTGAAGAGTTGTCCAAGAGCCCTCGACAAAGCCAGAGAGCCGACCTGCATT 3822
Qy 1147 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1166
Db 3823 GTGATTGCTCACCGCTGTCCACCATCCAGATCCAGACTTAATAGTGGTGTTCAGAAAT 3882
Qy 1167 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1186
Db 3883 GGCAGAGTCAAGAGAGTGGCAGCATCAGCAGCTGTGGCAGAGAAAGGCATCTATTTT 3942
Qy 1187 LysLeuValAsnAlaGlnSer 1193
Db 3943 TCAATGTCAGTGTCCAGGCT 3963

RESULT 6
5206352-3
Patent No. 5206352
Applicant: Koninson, Igor B.; Pastan Ira H.; Gottesman, Michael M.
Title of Invention: COMPOSITIONS FOR CLONES CONTAINING DNA
SEQUENCES ASSOCIATED WITH MULTIDRUG RESISTANCE IN HUMAN CELLS
Number of Sequences: 4
Current Application Data:
Application Number: US/07/622,836
Filing Date: 24-SEP-1990
Prior Application Data:
Application Number: 892,575
Filing Date: 01-AUG-1986
Application Number: 845,610
Filing Date: 28-MAR-1986
Seq ID No: 3
Length: 4669
5206352-3

Alignment Scores:
Pred. No.: 0 Length: 4669
Score: 3543.00 Matches: 681
Percent Similarity: 75.47% Conservative: 245
Best Local Similarity: 55.50% Mismatches: 265
Query Match: 58.93% Indels: 36
Gaps: 6
DB:

US-09-873-409-6 (1-1195) x 5206352-3 (1-4669)

Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 575 ATGGTGTGGGAATTGGCTGGCTCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 634
Qy 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
Db 635 GTGTTTGGAGAAATGACAGATATCTTTGCAAAATGCAGAAATTTAGAAAGATCTGATGCA 694
Qy 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
Db 695 AACATCTAATAGAGTATATCAATGATACAGGGTCTTCATGAATCTGGAGGAAGAC 754
Qy 44 -----ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61

Db 755 ATGACCAGGTATGCCTATTATTATACAGTGAATTTGGTGTGGGGTGTGGTGTGCTTAC 814
Qy 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db 815 ATTCAGGTTTCAATTTGGTGCCTGGCAGCTGGAAGCAAAATACACAAAATTTAGAAACAG 874
Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101
Db 875 TTTTTCATGCTATAATGCGACAGAGATAGGCTGTTTGATGTGCACGATGTGGGAG 934
Qy 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db 935 CTTAACACCCGACTTACAGATGATGTCTCAAGATTAATCAAGTTATTGGTGACAAAAT 994
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db 995 GGAATGTTCTTTTCAGTCAATGGCACTTTTTCACCTGGGTTTATAGTAGATTACACGT 1054
Qy 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db 1055 GGTGGAGCTAACCTTGTGATTGTCATCAGTCTCTGTTCTTGGACTGTCTAGCTGCT 1114
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db 1115 GTCTGGCAAGATACTATCTTCAATTTACTGATAAAGAACTCTTAGCTGATGCAAAAGCT 1174
Qy 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db 1175 GGACAGTAGCTGAGAGGTCTGGCAGCAATTAGAACTGTGATTGCTTGGAGACAA 1234
Qy 201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
Db 1235 AGAAGAACTTGAAGGTACACAAAATTTAGAAAGCTTAAAGAAATTTGGATAG 1294
Qy 221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
Db 1295 AAAGCTATTACAGCCCAATTTCTATAGTGTCTCTTCTCTGCTGATCTATGCTTAT 1354
Qy 241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
Db 1355 GCTCTGGCTTCTGGTATGGGACCACTTGGTCTCTCAGGGGAA-----TATTCTATT 1408
Qy 261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
Db 1409 GGACAGTACTACTGTATTCTTTCTGTTAATTTGGGGCTTTTAGTGTGACAGGCA 1468
Qy 281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
Db 1469 TCTCCAAGCATTTGAAGCAATTTGCAAAATGCAAGAGGACGCTTATGAAATCTTCAAGATA 1528
Qy 301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
Db 1529 ATTTGATAAATGAAGCAAGTATTGACGCTATTCGAAGTGGGCGACAAACAGATAATATT 1588
Qy 321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
Db 1589 AAGGGAATTTGGAATTCAGAAATGTTCACTTCACTTCCATCCATCTCGAAAGAGTTAAG 1648
Qy 341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
Db 1649 ATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGCGACAGCGTGGCCCTGGTTGGAAC 1708
Qy 361 AsnGlySerGlyLysSerThrValValGlnLeuGlnArgLeuTyrAspProAspAsp 380
Db 1709 AGTGGCTGTGGGAAGAGCACACAGCTCAGCTGATGAGAGGCTCTATGACCCACAGAG 1768
Qy 381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
Db 1769 GGAATGGTCAGTGTTCATGGACAGGATATTAGCACCAATAATGTAAGTTTCTACGGAA 1828
Qy 401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAla 420

Dbb 1829 ATCAATTGGTGGTGGAGTCCAGAACTCTGATTGTTGGCCACACGATAGCTGAACAACTT 1888
Qy 421 LysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
Dbb 1889 CGCTATGGCGGTGAAATGTCACCATGGATGAGATTGAGAAAGCTCTCAAGGAAGCCAAT 1948
Qy 441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
Dbb 1949 GCCTATGACTTTATCATGAACCTGCCTCATAAATTTGACACCCCTGGTGGAGAGAGGG 2008
Qy 461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Dbb 2009 GCCAGTTGAGTGGTGGCAGAAAGCAGAGATCGCATTTGCAGCTCCCTGGTTCGAAC 2068
Qy 481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Dbb 2069 CCCAAGATCCTCTCGTGGATGAGGCCACGTCAGCCTTGGACACAAAAGCGAAGCAGTG 2128
Qy 501 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 520
Dbb 2129 GTTCAGGTGGCTGTGATAAGGCCAGAAAAGGTTCGACCACCATTTGTGATGCTCATCGT 2188
Qy 521 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 540
Dbb 2189 TTGCTACAGTTTCGTAATGCTGACGTCATCGCTGGTTCGATGGAGTCATTTGTGGAG 2248
Qy 541 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 560
Dbb 2249 AAGGAAATCATGATGATCATCAAGAAAGAGAGGCAATTTACTTCAAACTTGTCACAATG 2308
Qy 561 Gln-----AspIleLysLysAlaAspGluGlnMetGluSerMetThr 574
Dbb 2309 CAGACAGCAGGAAATCAAGTTGAATTAGAAATGACGATGAATCCAAAGAGTGAAT 2368
Qy 575 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal----- 589
Dbb 2369 GATCCCTCGAATGCTCTCAATGATTCAGATCCAGTCCAGTCTAATAAGAAAAGATCAACT 2428
Qy 590 ---LysSerIleLys---SerAspPheIleAspLysAlaGluSerThrGlnSerLys 607
Dbb 2429 CGTAGAGGTGTCGTGGATCACAAAGCCCAAGACAGAAAGCTTAGTACCAAGAGGCTCTG 2488
Qy 608 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTyr 627
Dbb 2489 GATGAAAGTATACCTCCAGTTCTCTTGGAGGATATGAAGCTAAATTTAACTGAATGG 2548
Qy 628 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 647
Dbb 2549 CCTATATTTGTTGGTGTATTTTGTGCCATTATATAATGGAGGCTGCAACCCAGCATTT 2608
Qy 648 SerIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 666
Dbb 2609 GCNATATATTTCAAGATATATAGGGTTTTTACAGAAATGATGATCCTGAAACAAA 2668
Qy 667 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 686
Dbb 2669 CGACAGAATAGTAACCTGTTTCTCACTATTGTTCTAGCCCTTGGAAATATTTCTTTATT 2728
Qy 687 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 706
Dbb 2729 ACATTTTCTCTCAGGGTTTTCATATTGGCAAGCTGGAGAGATCCTCAACCAAGCGGCTC 2788
Qy 707 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaThrPheAspGluLysGlu 726
Dbb 2789 CGATACATGGTTTTCGATCCATGCTCAGACAGGATGTGAGTTGGTTGATGACCCCTAAA 2848
Qy 727 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 746
Dbb 2849 AACACCACTGGACATGATGATACAGGCTCGCCAAATGATGCTGCTCAAGTTAAGGGCT 2908
Qy 747 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 766
Dbb 2909 ATAGGTTCCAGGCTTGCTGTAAATTATCCAGAAATATAGCAAAATCTTGGACAGGAATAAT 2968

Qy 767 IleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeu 786
Dbb 2969 ATATCCTTTTCATCTATGTTGGCACTAACTACTTCTAGCAATTTGACCATTCATT 3028
Qy 787 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 806
Dbb 3029 GCAATAGCAGGAGTTGTTGAAATGTTGCTCGCAACAGCATGAAAGATAGAAA 3088
Qy 807 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 826
Dbb 3089 GAACTAGAAGGTGCTGGGAAGATCGTACTGAAGCAATAGAAAACCTTCGGAACCGTTGT 3148
Qy 827 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 846
Dbb 3149 TCTTTTACTCAGAGCAGAAAGTTTGAACATATGTATGTCTCAGAGTTTGCAGGTACCATAC 3208
Qy 847 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 866
Dbb 3209 AGAAACTCTTTGAGGAAGACACATCTTTGGAAATTACATTTTCTTACCAGGCAATG 3268
Qy 867 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 886
Dbb 3269 ATGTATTTTCTATGCTGGATGTTTCGGTTTGGAGCCTACTTGGTGCACATAAATC 3328
Qy 887 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 906
Dbb 3329 ATGAGCTTTGAGCATGTTCTGTAGTATTTTTCAGCTGTGTCTTGGTGCCTGCGGTG 3388
Qy 907 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaHisLeu 926
Dbb 3389 GGGCAAGTCAGTTCAATTTGCTCTGACTATGCAAAAGCCAAAATATCAGCAGCCCACT 3448
Qy 927 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluLysLysPro 946
Dbb 3449 ATCATGATCTTTGAAAAAACCCCTTGTGATGACGTACAGCACGGAAGCCCTAATGCCG 3508
Qy 947 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 966
Dbb 3509 AACACATTTGGAAGAAATGTCATTTGTTGTAAGTTGTATTCAACTATCCACCCGACCG 3568
Qy 967 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 986
Dbb 3569 GACATCCCACTGCTTCAGGGACTGAGCTTGGAGGTGAAGAGGGCCAGACGCTGGCTCTG 3628
Qy 987 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuGlnArgLeuTyrAsp 1006
Dbb 3629 GTGGCAGCAGTGGCTGTGGGAAGACACAGTGTCCAGCTCTCTGGAGCGGTCTACGAC 3688
Qy 1007 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTyr 1026
Dbb 3689 CCTTGGCAGGGAAGTGTCTGTATGTCAAAAGAAATAAAGCGACTGAATGTTCAAGTGG 3748
Qy 1027 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 1046
Dbb 3749 CTCGAGACACACTGGGCATCTGTGCCAGAGCCCATCTGTTTACCTGCAGCATTTGT 3808
Qy 1047 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 1066
Dbb 3809 GAGAACATTTGCTATGGAGAACACAGCCGGTGGTGTCAAGGAAGATCTGTGAGGGCA 3868
Qy 1067 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 1086
Dbb 3869 GCNAGGAGGCCAACATACATGCTTTCATCGAGTCACTGCCTAATAAATATAGCACTAAA 3928
Qy 1087 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 1106
Dbb 3929 GTAGGAGACAAAGAACTCAGCTCTCTGTGGTGGCCAGAAACAAACGCAATGCCATAGCTCGT 3988
Qy 1107 AlaLeuLeuGlnLysProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspAsn 1126
Dbb 3989 GCCCTTGTAGACGCTCATATTTTGTGTTTGGATGAAGCCACGTCAGCTCTGGATACA 4048

Db	2669	CGACAGATAGTAACCTGTTTTCACATATTGTTTCTAGCCCTTGGAAATTAATTTCTTTTATT	2712
Qy	687	SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluLeuLeuThrMetArgLeu	706
Db	2729	ACATTTTTTCTCCTCAGGGTTTCACATTTTGGCAAAAGCTGGAGAGATCCTCACCAGCGGCTC	2788
Qy	707	ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu	726
Db	2789	CGATACATGGTTTTTCGATCCCATGTCTCAGACAGGATGTGAGTTGGTTTTCATGACCCCTAAA	2848
Qy	727	AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla	746
Db	2849	AAACACATGGACCATTGACTACAGGCTCGCCAAATGATGCTGCTCAAGTTTAAAGGGCT	2908
Qy	747	ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle	766
Db	2909	ATAGGTTCCAGGCTGCTCTGTAATTTACCCAGAAATATAGCAAATCTTGGGACAGGAATAATT	2968
Qy	767	IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu	786
Db	2969	ATATCCTTCATCTATGTTTGGCAACTTAACACTGTTTACTCTTTAGCAATTTGTACCCATCATT	3028
Qy	787	AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln	806
Db	3029	GCAATAGCAGGAGTTGTTGAAATGAAATGTTTGTGTCGACAGACNCTGAAGATATAGAAA	3088
Qy	807	GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal	826
Db	3089	GAACTAGAGGTCCTGGGAAGATCGCTACTGAAGCAATAGAAAAACTTCCGAACCGTTGTT	3148
Qy	827	SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis	846
Db	3149	TCCTTGACTCAGGACGAGAAGTTTGAACATATGTATGCTCAGAGTTTCCAGGTACCATAC	3208
Qy	847	ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe	866
Db	3209	AGAACTCTTTGGAGGAACACACATCTTTGGAATTAACATTTTCTTCACCAGCGCAATG	3268
Qy	867	IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg	886
Db	3269	ATGTAATTTTCTATCTCGATGTTTCCGGTTTGGAGCTACTTGGTGGCACATAAACTC	3328
Qy	887	MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle	906
Db	3329	ATGAGCTTTTGAGGATGCTCTGTAGTATTTTCAGCTGTTTCTTTTGGTGCATGCCGCTG	3388
Qy	907	GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu	926
Db	3389	GGGCAAGTCAGTTTCATTTGCTCTGCTACTATGCCAAGCCAAATATCATGAGCCCCACAT	3448
Qy	927	PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLeuPro	946
Db	3449	ATCATGATCATTGAAAAACCCCTTTGATTGACAGCTACAGCACGGAGGCTTAATCGC	3508
Qy	947	AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro	966
Db	3509	AACACATTGGAAGAAATGTCACTTTGTTGAAGTTGTATTCAACTATCCACCCCGCCG	3568
Qy	967	AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe	986
Db	3569	GACATCCAGTGCTTCAGGACATGAGCCTGGAGGTGAAGAAGGGCCAGACCGCTGCTCTG	3628
Qy	987	ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp	1006
Db	3629	GTGGGCAGCAGTGGCTGTGGGAAGACACAGTGGTCCAGCTCCTGGAGCGGTTCTACGAC	3688
Qy	1007	ProValGlnGlnGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp	1026
Db	3689	CCCTTGGCAGGAAAGTGCTGCTTGATGCCAAAGAAATAAAGCCGACTCAATGTTTCAGTGG	3748
Qy	1027	LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla	1046
Db	3749	CTCGAGCACCACTGGGCATCGTGTCCAGAGAGCCCATCTCTGTTTGCATGACGATTGCT	3808

Db	2369	GATGCTCTGGAAATGTCTTCAAATGATTCAGATCCAGTCTTAATAGAAAAAGATCAACT	2422
Qy	590	---LysSerIleLysSerAspPheIleAspLysAlaGluGluSerThr---GlnSerLys	607
Db	2429	CGTAGGAGTGTCCGTGGATCACAAGCCCAACACAGAAAGCTTAGTACCAAAAGGCTCTG	2488
Qy	608	GluLeSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTirp	627
Db	2489	GATGAAATAGTACTCCAGTTTCCTTTTGGAGGATATGAAGCTAAATTTAACTGMAATGG	2548
Qy	628	ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe	647
Db	2549	CCTTATTTTGTGTGGTGATTTTGTGCGCATTAATAATGAGGCTCGAACACGACATTT	2608
Qy	648	SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu	666
Db	2609	GCAATAATATTTCAAAGATTATAGGGTTTTTCACAGAAATTGATGATCCTGAAACAAA	2668
Qy	667	LysHisAspAlaGluLeuTyrSerMetIlePheValIleLeuGlyValIleCysPheVal	686
Db	2669	CGACAGAAATAGTAACCTGTTTTCACATATTGTTTCTAGGCCCTTGGAAATATTCTTTTATT	2728
Qy	687	SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu	706
Db	2729	ACATTTTTTCTCTCAGGGTTCACATTTGGCAAAAGCTGAGAGATCCTCACCAGCGGCT	2788
Qy	707	ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTirpPheAspGluLysGlu	726
Db	2789	CGATACATGGTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTCATGACCCATAA	2848
Qy	727	AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla	746
Db	2849	AACACCACTGGAGCAATTGACTACACGGCTCGCCAATGATGCTGCTCAAGTTTAAAGGGCT	2908
Qy	747	ThrClySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetClyLeuSerValIle	765
Db	2909	ATAGGTTCCAGGCTGCTGTAATATCCCAAGATATAGCAAAATCTTGGGACAGGAATAAT	2968
Qy	767	IleSerPheIleTyrGlyTirpGluMetThrPheLeuIleLeuSerIleAlaProValLeu	786
Db	2969	ATATCCTTCATCTATGGTTGGCACTAAACACTGTTACTCTTAGCAATTTGATCCCATCAT	3028
Qy	787	AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln	806
Db	3029	GCAATAGCAGGAGTTGTGAAATGAAATGTTTGTGTGGCAAGCACTGAAAGATAAGAAA	3088
Qy	807	GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal	826
Db	3089	GAACTAGAAGGCTCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCGGAACCGTGTGT	3148
Qy	827	SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluMetLeuGlnThrGlnHis	846
Db	3149	TCTTTTGACTCAGAGCAGAAGTTTGHACATATGTATGCTCAGAGTTTCAGGTACCATAC	3208
Qy	847	ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe	866
Db	3209	AGAAACTCTTTGAGGAAAGCACACATCTTTTGGAAATTACATTTTCTCACCACCGCAATG	3268
Qy	867	IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg	886
Db	3269	ATGTATTTTTTCTATGCTGGATGTTTTCCGGTTTGGAGGCTACTTGGTGGCACAATAAACTC	3328
Qy	887	MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle	906
Db	3329	ATGAGCTTTCAGGATGTTCTGTGTAGTATTTTTCAGCTGTGTGCTTTGGTGCCATGCCGTG	3388
Qy	907	GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaHisLeu	926
Db	3389	GGGCAAGTCAGTTTCATTTGCTCTCTGATGATGCCAAAGCCAAAAATATCAGCAGCCACATC	3448
Qy	927	PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro	946
Db	3449	ATCATGATCATGAAAAAACCCCTTTGATTGATGACGTACAGCAGGAAGGCTAATGCCG	3508


```
QY 947 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 966
Db 3509 AACACATTGGAAGAAATGTCACATTGGTGAAGTTGTTCAACTATCCACCCGACCG 3568
QY 967 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 986
Db 3569 GACATCCAGCTGTTGAGGACTGAGCTGGAGTGAAGAGGCCAGACGCTGGCTCTG 3628
QY 987 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuGluArgLeuTyrAsp 1006
Db 3629 GTGGGACGAGTGGCTGGGAAGACACAGTGTCCAGCTCTCTGGAGCGGTCTCAGC 3688
QY 1007 ProValGlnGlyGlnValLeuPheAspGlyValAlaLysGluLeuAsnValGlnTyr 1026
Db 3689 CCCTGGCAGGAAAGTCTGCTGTGATGGCAAGAAATAAGCGACTGAATGTTCACTG 3748
QY 1027 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 1046
Db 3749 CTCGAGCACACCTGGGCATCGTGTCCAGGAGCCCTCTGTTGACTGCAGCATTCGT 3808
QY 1047 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 1066
Db 3809 GAGAACATTGCTATGAGACACAGCCGGGTGTGTCAAGGAGAGATCGTGGGGCA 3868
QY 1067 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 1086
Db 3869 GCAAGAGGAGCCACATACATGCTTCATCGACTGCTAATAATAATAGCACTAAA 3928
QY 1087 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 1106
Db 3929 GTAGGACAAAGAACTGAGTCTCTGTGTGGCCAGAAACAACGCAATGCGATGCTCGT 3988
QY 1107 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 1126
Db 3989 CGCCTTGTAGACAGCCTCATATTTGCTTTGGATGAAGCCAGCTCAGCTCTGGATACA 4048
QY 1127 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1146
Db 4049 GAAAGTGAAGAGTTGTCCAAAGAGCCCTGGACAAAGCCAGAGAGCCGCACTGCATT 4108
QY 1147 ValValThrHisArgSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1166
Db 4109 GTGATTCTCACCGCTGTCCACCATCCAGATGCGAGCTTAATAGTGTGTTCAGAAAT 4168
QY 1167 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1186
Db 4169 GGCAGAGTCAAGGAGCATGGCAGCGCATCGACGCTGTGGCACAGAAAGGCATCTATTTT 4228
QY 1187 LysLeuValAsnAlaGlnSer 1193
Db 4229 TCAATGTCAGTGTCCAGGCT 4249

RESULT 9
US-09-120-513-1
; Sequence 1, Application US/09120513
; Patent No. 6025160
; GENERAL INFORMATION:
; APPLICANT: Brun, Kimberly
; APPLICANT: Chenery, Richard
; APPLICANT: Ellens, Harma
; APPLICANT: Field, John
; APPLICANT: Yue, Lin
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE
; TITLE OF INVENTION: SEQUENCES ENCODING RAT MDR1B2 AND
; TITLE OF INVENTION: SCREENING METHODS THEREOF
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY:
```

```
ZIP: 19406
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/120.513
FILING DATE: 22-JUL-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: King, William T
REGISTRATION NUMBER: 30,954
REFERENCE/DOCKET NUMBER: GP50008
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5015
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4233 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
US-09-120-513-1

Alignment Scores:
Pred. No.: 0 Length: 4233
Score: 3517.50 Matches: 684
Percent Similarity: 74.47% Conservative: 229
Best Local Similarity: 55.79% Mismatches: 274
Query Match: 58.51% Indels: 39
DB: 3 Gaps: 9

US-09-873-409-6 (1-1195) x US-09-120-513-1 (1-4233)
QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyValAlaCysLeuProLeuMetProLeu 20
Db 170 ATGGCTCTGGGAACCTCTCGTGTATCATCCACGGAACCTCTCCCTCCCTGTGATGCTG 229
QY 21 ValLeuGlyGluMetSerAspAsnLeuIle----- 30
Db 230 GTGTTCCGATATACATGACAGATAGTTTACCCAGCAGACAGCCGCAATTCCTGCGAGCGTT 289
QY 31 -----SerGlyCysLeuValGlnThrAsnThrTyrSerPhePheArgLeu----- 45
Db 290 ACTAATCAAAGTGAATCAACACAGTACACAGACCGTCAGCAGACAGCTCTGGAGGAGGAC 349
QY 46 -----ThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db 350 ATGGCCATGTACGCTACTATTACAGCGGCATTTGGTCCGCTGTGCTCATCGTTGCTTAC 409
QY 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db 410 ATCCAGGTTTCACTTTGGTGCCTGGCAGCTGGGAGACAAATACACAGATTAGGCAGAAG 469
QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101
Db 470 TTTTTCATGCCATCATGAATCAGGAGATAGGCTGTTGACGTGAATGACGCTGGGGAG 529
QY 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db 530 CTCACACCCGGCTCACATGACGTCTCCAAATTAATGACGGAATTTGGTGAACAATTT 589
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db 590 GGAATGTTCTTTTCAGTCCATAACGACATTTTCAGCCGCTTTTATATAGGATTTAAGT 649
QY 141 GlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db 650 GGTTGGAAGCTAACCCCTTGTAAATTTTGGCCGCTCAGCCCTCTTATGGTGTGTCATCTGCC 709
```


Qy	161	AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla	180
Db	710	ATGTGGCCAAAGGTACTGACTTCATTTACTAATAAGGAACCTCCAGGCTTATGGCAAGCT	769
Qy	181	GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln	200
Db	770	GGACAGTTCGCCAAGAAGTCTTTAGCAGCCATCAGAACTGTGATGCTTTGGAGAGCAA	829
Qy	201	GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys	220
Db	830	AAGAAGGAACCTTGAAGGTACATAAANAATTTAGGAAGACTAAAGAGTTGGCATAAAG	889
Qy	221	ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr	240
Db	890	AAAGCCATCAGGCCAACATTTCCATAGTATTGGCTACCTGTTGGTCTCATGCGCTCTAT	949
Qy	241	GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle	260
Db	950	GCATCGCATCTCGTATGGAGCTCTCTGGTGCCTCTCAATGAATGAA-----TATCTATT	1003
Qy	261	GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla	280
Db	1004	GGACAAGTGCTTACCGCTCTCTCTCTATTATTATGGGGACTTTCAGTATTGGACATTTA	1063
Qy	281	ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal	300
Db	1064	CCCCAAACATAGAAGCCTTTGCNAATCAAGAGGGCGAGCTTATGAAATCTTCAAGATA	1123
Qy	301	IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle	320
Db	1124	ATTGATTAATGAGCCACGACATGTGACGCTTCTCAACCAAGGGACACAACACGACAGTATA	1183
Qy	321	GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys	340
Db	1184	ATGGGAATTTGGNAATTTAAATGTTTACTTCACTACCCATCAGGAAGTGAAGTTAAG	1243
Qy	341	IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu	360
Db	1244	ATCTTGAAGGCCTCAACCTGAAGGTGAAGAGCGGCGACGCGTAGCCCTGGTTGGCAAC	1303
Qy	361	AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp	380
Db	1304	AGTGGCTGTGGGAAAAACCAACTGTCCAGCTGCTGCAGAGGCTCTACGACCCCATGAG	1363
Qy	381	GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp	400
Db	1364	GGCAGGTCAGTATCGACGACGACGACATCAGGACCATCAATGTAGGTATCTCGGGAA	1423
Qy	401	HisIleGlyValValSerGlnGluProValLeuPheGlyThrIleSerAsnAsnIle	420
Db	1424	ATCATTGGGGTGTGAGTCAGGAACCGTGTGTTTGGCCACCACGATTTGCCGAAACATT	1483
Qy	421	LysTyrGlyArgAspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsn	440
Db	1484	CGCTATGGCCGAGAAAAACGTCCACATGGATAGATAGAAAAAGCTGTCAAGGAAGCCCAAT	1543
Qy	441	AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGly	460
Db	1544	GCCTATGACTTCATGAACATGCTCCCCACAAATTTAACACCTGGTGGTGAGAGAGGG	1603
Qy	461	AlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn	480
Db	1604	CGCAGCTGAGTGGGGACAGAAACAGAGGATCGCCATTTGCCCGGCCCTGGTCGCAAC	1663
Qy	481	ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla	500
Db	1664	CCCAAGATCTCTTTGTTGGATAGGGCCACGTGACGCTCGACACAGAAGCGAAGCCGCTG	1723
Qy	501	ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg	520
Db	1724	GTTTCAGGCCGCTCTGGATATAGGCTAGAGAGCCGACCAACCATTTGTGATAGTCTCACCGC	1783
Qy	521	LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu	540

[illegible]

1904	Db	CAG---ACTAGAGAAATGAAATTTGACACAGAAATAATGCTTATGAATCCCAAGTGCAC	1960
581	Qy	ThrAsnSerLeuProLeuHisSerValIysSerIleIysSerAspPheIleAspIysAla	600
1961	Db	ACTGGTGCCTCTGAGTTGACTTCAGAAAAATCA---AAATCTCCTTTAATAAGGAGTCA	2017
601	Qy	-----GluGluSerThrGlnSerIysGlu---	608
2018	Db	ATTGCGCAGAGTATATCACAGAAAGACCAAGACAGGAGAGAGACTTAGTTTCGAAAGAGAT	2077
609	Qy	-----IleSerLeuProGluValSerLeuLeuIysIleLeuIysLeuAsnIysProGlu	626
2078	Db	GTGGATCAAGATGTGCTATGTTTCCTTTTGGCAGATCCTTAAGCTAAATATATAGTGAA	2137
627	Qy	TrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProVal	646
2138	Db	TGGCCCTATTAGTTGTGGGTGACTTTGTGCTGTATATAAATGGGTGCATACCAACGAGT	2199
647	Qy	PheSerIleIlePheAlaIysIleIleThrMetPheGly--AsnAsnAspIysThrThr	665
2198	Db	TTTGCCATAGTGTTTTCAAGATGTGTAGGGGTTTTTCAAGAGACGACGACCATGAACCC	2257
666	Qy	LeuIysHisAspAlaGluIleTySerMetIlePheValIleLeuGlyValIleCysPhe	685
2258	Db	AAACACCGAATGTAACTTGTTTCCCTTCCTTCTTGCTCATGGGAATGATTTCTTTT	2317
686	Qy	ValSerTyrrPheMetGlnGlyLeuPheTyrrGlyrArgAlaGlyGluIleLeuThrMetArg	705
2318	Db	GTTCAGGTACTTCCTTCAAGGCTTCACATTTTGCAAGCTGGAGAGATCCTACCAACGCA	2377
706	Qy	LeuArgHisLeuAlaPheIysAlaMetLeuTyrrGlnAspIleAlaTrpPheAspGluIys	725
2378	Db	CTCGATACATGCTCTCAAAATCCATGCTCGCAGCAGGATATAAGCTGGTTTGATGACCAT	2433
726	Qy	GluAsnSerThrGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGly	745
2438	Db	AAAAACACCACTGGCTCGCTGACTACGAGCTCGCTAGTGAGCGCTTCTAAGTTTAAAGGG	2497
746	Qy	AlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerVal	765
2498	Db	GCTATGGGCTCCAGGCTTGCTGTAGTTATCCAGAAATGTAGCAAACTTGGGCACAGGAAT	2555
766	Qy	IleIleSerPheIleTyrrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProVal	785
2558	Db	ATCTTATCCTTAGTCTATGGCTGGCAGCTTACACTTTTACTTTGTAGTAATATATACCATC	2611
786	Qy	LeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnIysAspIys	805
2618	Db	ATTGCTTTGGGTGGAAATATTGAATGAACACTGTTGTCTGGTCAAGCCTTGANGACGAG	2677
806	Qy	GlnGluLeuIysHisAlaGlyIysIleAlaThrGluAlaLeuGluAsnIleArgThrIle	825
2678	Db	AAAGAGCTAGAGATCTCTGGGAAGATCGCTACAGAAGCAATTTGAAAACTTCCGCACATGT	2733
826	Qy	ValSerLeuThrArgGluIysAlaPheGluGlnMetTyrrGluGluMetLeuGlnThrGln	845
2738	Db	GTCTCTTTGACTAGGAGCAGAGAAGTTTGAACAATACTATGTATGCCACAGAGCTTGCA	2799
846	Qy	HisArgAsnThrSerIysIysAlaGlnIleIleGlySerCysTyrrAlaPheSerHisAla	865
2798	Db	TACAGAAATGCTTTGAAGAAGACACAGCTCTTTGGGATCACCTTCGCTTTCACCCAGGCC	2855
866	Qy	PheIleTyrrPheAlaTyrrAlaAlaGlyPheArgPheGlyAlaTyrrLeuIleGlnAlaGly	885
2858	Db	ATGATTTATTTTTCCTATGCTGCTTGTTCCTGGTTCGGTTCGGCTACTTGGTGGCACGAG	2911
886	Qy	ArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrrGlyAlaMetAla	905
2918	Db	CTCATGACGTTTGAAAATGTTATGTGTGGTATTCTTGCTGTGTGTCTTGTGGTGCCTATG	2977
906	Qy	IleGlyIysThrLeuValLeuAlaProGluTyrrSerIysAlaIysSerGlyAlaAlaHis	925
2978	Db	GCAGGGAAATACCAAGTTTCATTCGCTCTCTGACTACGGAGAGGCCAAAGTCTCGATCCTCC	3037

QY 926 LeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLys 945
Db 3038 ATCATCAGATCATTTAGAAATATCCCGAGATTGACAGCTACAGCAGGAGGCTTGAAG 3097
QY 946 ProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArg 965
Db 3098 CCTAATGGTTAGAAAGAAATGTAATTTAATGGAGTCATGTTCAACTATCCACCCGA 3157
QY 966 ProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAla 985
Db 3158 CCAACATCCAGTGTTCAGGAGCTGAGCTTCGAGTGAAGAGGGGCAACCGCTTCGC 3217
QY 986 PheValGlySerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyr 1005
Db 3218 CTGGTGGCAGCAGTGGCTCGGAGAGAGTACAGTGGTCAGCTGCTCAGAGCTTCTAC 3277
QY 1006 AspProValGlnGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGln 1025
Db 3278 AACCCATGGCTGGAACAGTGTTCAGTGGCAAGAAATAAACAACCTCAACGCTCCAG 3337
QY 1026 TrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIle 1045
Db 3338 TGCTCCGCCGCCACCTGGGCATGTGTCCAGAGGCCCATCTCTGTGACTCGAGCATC 3397
QY 1046 AlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGlu 1065
Db 3398 ACCGAGACATCCCTACGAGAGCAACAGCCGTGTCTGTCTCATGAGGAGATCGTGAGG 3457
QY 1066 AlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThr 1085
Db 3458 GCCGCCAGGAGGCCAACATCCACCACTTCATCGACTCATCTGCTGAGAAATACAACACC 3517
QY 1086 GlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAla 1105
Db 3518 AGAGTGGGACAAAGGAGTCACTGTCTCGGGGGGAGAGCGCATCGCCATCGCGC 3577
QY 1106 ArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAsp 1125
Db 3578 CGGCCCTCTGACAGACCTCACTCACTTACTTCTGATGAGGAGCATCAGCTCTGGAT 3637
QY 1126 AsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgCys 1145
Db 3638 ACGGAGTCAAAAGGTCTGCCAGGAGCGCTGGACAAAGCCAGGAGGCCCGCATCTGC 3697
QY 1146 LeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHis 1165
Db 3698 ATTGTGATCGGCACCGCTGTCCACCATCCAGAACGACACTTGATCGGTGATTTCAG 3757
QY 1166 AsnGlyLysIleLysGluGlnGlyThrHisGlnGlnLeuLeuArgAsnArgAspIleTyr 1185
Db 3758 AAGGCCAGTCAAGGAGCAGCGGCCACCCACACAGCTGTCTGGCCCAAGAGCATCTAT 3817
QY 1186 PheLysLeuValAsnAla 1191
Db 3818 TTCTCGATGGTTGAGGCT 3835

RESULT 11

US-08-583-276-18

Sequence 18, Application US/08583276

Patent No. 5837536

GENERAL INFORMATION:

APPLICANT: McDonagh, Kevin T.

APPLICANT: Nienhuis, Arthur

APPLICANT: Tolstoshev, Paul

TITLE OF INVENTION: IMPROVED EXPRESSION OF HUMAN

TITLE OF INVENTION: MULTIDRUG RESISTANCE GENES AND IMPROVED

TITLE OF INVENTION: SELECTION OF CELLS TRANSFECTED WITH SUCH GENES

NUMBER OF SEQUENCES: 19

CORRESPONDENCE ADDRESS:

ADDRESS: Carella, Byrne, Bain, Gilfillan,

ADDRESS: Cecchi & Stewart

STREET: 6 Becker Farm Road

CITY: Roseland
STATE: New Jersey
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch diskette
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: DM4 V2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/583,276
FILING DATE: 05-JAN-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/332,444
FILING DATE: 31-OCT-1994
APPLICATION NUMBER: 07/887,712
FILING DATE: 22-MAY-1992
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 4669 bases
TYPE: nucleic acid
STRANDEDNESS: singular
TOPOLOGY: linear
MOLECULE TYPE:
DESCRIPTION: Genomic DNA
US-08-583-276-18

Alignment Scores:

Pred. No.: 0 Length: 4669
Score: 3497.00 Matches: 673
Percent Similarity: 74.96% Conservative: 246
Best Local Similarity: 54.89% Mismatches: 271
Query Match: 58.17% Indels: 36
DB: 2 Gaps: 9

US-09-873-409-6 (1-1195) x US-08-583-276-18 (1-4669)

QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyValaCysLeuProLeuMetProLeu 20
Db 575 ATGTGGTGGGAACCTTGGCTGCCATCATCGGGCTGGAGCTTCTCTCATGATGCTG 634
QY 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
Db 635 GTGTTTGGAGAAATGACAGATATCTTTGCAATGCAGGAAATTTAGACAGATCTGATGTC 694
QY 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
Db 695 AACATCACTAATAGAGTATCATATCATGAGGCTTCTTCATGAATCTGGAGGAGAC 754
QY 44 -----ArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db 755 ATGACACAGTATGCCATTATTATACAGTGAATGGCTGGGGTGTGGTGTGCTGTAC 814
QY 62 IleGlnIleSerLeuTrpIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db 815 ATTCAAGTTTCAATTTGGTGGCTGGCAGCTGGAAGACAAATACACAAATTTAGAAACAG 874
QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101
Db 875 TTTTTCATGCTATATGCGCAGGAGATAGGCTGTTGATGTGCACGATGTTGGGCGC 934
QY 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db 935 CTTAACACCCGACTTACAGATGATGCTCTTAAGATTAAATGAATTTATGTTGACAAAT 994
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db 995 GGAATGCTCTTTCAGTCAATGGCAACATTTTCACCTGGGTTTATAGTAGGATTTACAG 1054
QY 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db 1055 GGTGGAAGCTAACCTTGTGATTTTGGCCATCATGCTCTGTTCTTGGACTGTGAGTGT 1114

Qy	161	AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla	180
Db	1115	GTCTGGGCAAAAGTACTATCTTCATTACTGATAAAGAACTCTTAGCGTATGC AAAAGCT	1174
Qy	181	GlyAlaValaIaGluValLeuSerSerLysArgThrValIleAlaPheArgAlaGln	200
Db	1175	GGAGCAGTAGCTCAAGAGGCTTGGCGCAATTTAGAAGCTGTGATTGCAATTTGGAGGACAA	1234
Qy	201	GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys	220
Db	1235	AAGAAAGAACTTGAAGAGTACAACAAAATTTAGAAGAGCTTAAAGAAATTTGGGATTAAG	1294
Qy	221	ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr	240
Db	1295	AAAGCTATTACAGCCAATATTTCTATAGTGTCTCTTCCTGCTGATCTATGCATCTTAT	1354
Qy	241	GlyLeuAlaPheTyrPyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle	260
Db	1355	GCTCTGGCCCTCTGGGTATGGGACCACTTGGCTCTCTCAGGGAA-----TATTTTATT	1408
Qy	261	GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla	280
Db	1409	GGACAAGTACTCACTGTATCTTTTCTGTATTAATTTGGGGCTTTTAGTGTGGACAGCA	1468
Qy	281	ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal	300
Db	1469	TCCTCAAGCATTTGAAGCAATTTGCAAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA	1528
Qy	301	IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle	320
Db	1529	ATTGATAATTAAGCCAAGTATTGACAGTATTCGAAGAGTGGGCGACAAACAGACATAATTT	1588
Qy	321	GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys	340
Db	1589	AAGGGAATTTGGAATTCGAAATGTTCACTTCAGTATACCATCTCGAAAGAAAGTTTAAG	1648
Qy	341	IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu	360
Db	1649	ATCTTGAGGCGCTGAACCTGAAGTGACAGTGGGCGACAGTCCCTTGGTTGGAAAC	1708
Qy	361	AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp	380
Db	1709	AGTGGCTGTGGGAAGAGCACAAACAGTCCAGCTGATGTCAGAGGCTCTATGACCCACACAG	1768
Qy	381	GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp	400
Db	1769	GGGATGGTCACTGTTTCATGCACAGGATATTTAGGACCATTAATGTAAAGGTTTCTACGGGAA	1828
Qy	401	HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle	420
Db	1829	ATCATTTGGTGTGTAGTCAGGAACCTGTATTTGTTGGCCACCATGATCTGTAAGAAACATT	1888
Qy	421	LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn	440
Db	1889	CGCTATGGCGCTGAAAAATGTCCACATGGATGAGATTTGAGAAGGCTGTCAAGGAAGCCAA	1948
Qy	441	AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly	460
Db	1949	GCCTATGACTTTATCATGAACCTGCCTCATAAATTTGACACCCCTGGTTGGAGAGAGAGG	2008
Qy	461	AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn	480
Db	2009	CCCCAGTTTGAAGTGGTGGCGAAGACAGAGATGGCCATTTGCACGTGCCCTGTTTCGCAAC	2068
Qy	481	ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla	500
Db	2069	CCCAAGATCTCTCTGCTGGATGAGGCCAGCTCAGCTTGGACACAGAAAGGCAAGCAGTG	2128
Qy	501	ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg	520
Db	2129	GTTCCAGTGGCTCTCGGATAAGGCCAGAAAGGTGGACCAACCATTTGTGATAGCTCATCGT	2188

Qy	521	LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu	540
Db	2189	TTGTCTACAGTTCGTAAATGTCACGCTTCGATGCGAGTCTGATGCG	2248
Qy	541	LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrSerLeuValMetSer	560
Db	2249	AAAGGAATCATGATGAACATCATGAAGAAGAAGCAATTACTTCAAACTTGTCCAAATG	2308
Qy	561	Gln-----AspIleLysLysAlaAspGluGlnMetGluSerMetThr	574
Db	2309	CAGACACAGGAATGAAGTTGAATTAGAAAATGCAGCTGATGATCCAAAAGTGAAATT	2368
Qy	575	TyrSerThrGluArgLysThrIleSerLeuProLeuHisSerVal-----	589
Db	2369	GATCCCTTGGAAATGCTTCAAAATGATCAAGATCACGCTCAATTAAGAAAAAGATCAACT	2428
Qy	590	--LysSerIleLys--SerAspPheIleAspLysAlaGluSerThrGlnSerLys	607
Db	2429	CGTAGGAGTGTGCTGGATCACACCCCAAGACAGAAGACTTAGTACCAGAAGGCTCTG	2488
Qy	608	GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTyr	627
Db	2489	GATGAAGTAGTACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGG	2548
Qy	628	ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe	647
Db	2549	CCTTATTTCGTGTGTGTATTTTGTGCCATATAAAATGGAGGCTGCCAACACGACATTT	2608
Qy	648	SerIleIlePheAlaLysIleIleThrMetPheGlyAsn--AsnAspLysThrThrLeu	666
Db	2609	GCAATAATATTTCAAAGATTATAGGGGTTTTTACAAGAATTGATGATCCTGGAACAACAAA	2668
Qy	667	LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal	686
Db	2669	CGACAGAACTACTAATCTGTTTTCACATATGTTTCTAGCCCTTGGAAATTAATTTCTTTATT	2728
Qy	687	SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu	706
Db	2729	ACATTTTTCCTTCAGGCTTTCACATTTGGCAAAAGCTGGAGAGATCCTCACCACGCGCTC	2788
Qy	707	ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGluLysGlu	726
Db	2789	CGATACATGTTTTCGATCCATGCTCAGACAGAGATGAGTGTGTTGATGACCCCTAAA	2848
Qy	727	AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla	746
Db	2849	AACACCACTGGAGCATTGACTACCAGGCTCGCAATGATGCTGCTCAAGTTAAAGGGCT	2908
Qy	747	ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle	766
Db	2909	ATAGTTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAATTTCTTTGGGACAGGAATAAT	2968
Qy	767	IleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeu	786
Db	2969	ATATCCTTCATCTATGTTGGCAACTAAACACTGTACTCTTAGCAATTTGTACCACATCAAT	3028
Qy	787	AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln	806
Db	3029	CAATACAGAGATTGTTGNAATGAATAATGTTGTTCTGGACAGCACTGGAAGATAGAATA	3088
Qy	807	GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal	826
Db	3089	GAACTAGAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTGTT	3148
Qy	827	SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis	846
Db	3149	TCCTTTGACTCAGGACGAGAAGTTTGAACATATGTTATGCTCAGAGTTTTCAGCGTACCATAC	3208
Qy	847	ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe	866
Db	3209	AGAAACTCTTTGAGGAAGACACACATCTTTGGAATTAACATTTCTCTCACCAGGCAATG	3268
Qy	867	IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg	886

Db	3269	ATGTAATTTTTCCTATCTGGATGTTTCCGGTTTGAGCGCTACTTGGTGGCACAATAAACTA	3328
Qy	887	MetThrProGluGlyMetPheIleValPheThrAlaIleAlaIleValGlyAlaMetAlaIle	906
Db	3329	ATGAGCTTTGAGGATGCTCTGTTAGTATATTTTTCAGCTGTGTTCTTTGGTGCCATGGCCGTG	3388
Qy	907	GlyIysThrLeuValLeuAlaProGluTyrSerIysAlaIysSerGlyAlaAlaHisLeu	926
Db	3389	GGGCAAGTGAGTTTCATTTGCTCTGACTATGCCAAGGCCAAATATATCAGACGCCACATC	3448
Qy	927	PheAlaLeuLeuGluIysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro	946
Db	3449	ATCATGATCATTTGAAAAACCCCTTTTGATTGACAGCTACAGCACGGAGGCGCTAATCCG	3508
Qy	947	AspThrCysGluGlyAsnLeuGluPheArgGluValSerPheTyrProCysArgPro	966
Db	3509	AACACATTTGGAAGGAATGTCACTTTGTTGTAAGTTGTATTTCAACTATCCACCCGACCG	3568
Qy	967	AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe	986
Db	3569	GACATCCAGTCTCTCAGGAGCTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTG	3628
Qy	987	ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp	1006
Db	3629	GTGGCAGCAGTGGCTGTGGGAAGACACAGATGGTCCAGCTCTCTGGAGCGGTTCTACGAC	3688
Qy	1007	ProValGlnGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTyr	1026
Db	3689	CCCTTGGCAGGGAAGTGGCTTGTATGCCAAGAAATAAAGCAGCATGAATTTTCAGTGG	3748
Qy	1027	LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla	1046
Db	3749	CTCGAGCACACTGGGCATGTCGTCCAGAGAGCCCATCTCTGTGTGATGTCGAGCATTGCT	3808
Qy	1047	GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla	1066
Db	3809	GAGAACTTCCTATGGACACACACAGCCGGTGGTGTCCACAGGAAGAGATCGTGAGGCGCA	3868
Qy	1067	AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln	1086
Db	3869	GCAAAGAGGCCAACATACATGCTTCATCGAGTCACTGCTCTTAATATATAGCACTAAA	3928
Qy	1087	ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg	1106
Db	3929	GTAGGACACAAAGCACTCAGCTCTCTGGTGCCAGAAACAACGCAATGCCATAGCTCGT	3988
Qy	1107	AlaLeuLeuGlnLysProIysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn	1126
Db	3989	GCCCTGTTAGACAGCCTCATATTTTGGCTTTTGGATGAAGCCAGCCAGCTCTGGATACA	4048
Qy	1127	AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu	1146
Db	4049	GAAGTGTAAAAAGTTGTCCAAAGAGCCCTTGGACAAACCCAGAGAGCGCGCACTGCATT	4108
Qy	1147	ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn	1166
Db	4109	GTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTCAGAAT	4168
Qy	1167	GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe	1186
Db	4169	GGCAGATCAAGCAGATGGCAGCATCAGCATGCTGCTGGCAGAGAAGGCATCTATTTT	4228
Qy	1187	LysLeuValAsnAlaGln	1192
Db	4229	TCAATGGTCAGTGCCAG	4246

RESULT 12

US-08-996-545-1

00 00 000 010 1
: Sequence 1, Application US/08996545

: Patent No. 5928898

; FALCILE NO. 5928858
: GENERAL INFORMATION:

APPLICANT: skatrud. paul L.

[illegible][illegible]

Db 2803 CAGTTTCAATCAACGCTCCAAAGCTTGCTATGAGGGATGTGCAAACTTTGCTTCGCGAGCT 2862
Qy 819 LeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyr 838
Db 2863 ACATCGTCTATCCGACAGTGTCTATTAACCCGGAAGGATGTCTGGGAGATTTC 2922
Qy 839 GluGluMetLeuGlnThrHisArgAsnThrSerLysLysAlaGlnIleIleGlySer 858
Db 2923 CATGCCAGCTTGACGCACACAGCAGCAGCAGCTAATCTCTCTGTGAGGTATCCCTG 2982
Qy 859 CysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaIleGlyPheArgPheGly 878
Db 2983 TTATATGCGTGTGCGCAGCAGCTGTCTTCTCTCGCTGCGCTGCGGTCTTGTACGGA 3042
Qy 879 AlaTyrIleuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAla 898
Db 3043 GGCACACTTCTTGCTACACAGATATGACATTTTCGCTTCTTTGTTGTTCTCTCCGAG 3102
Qy 899 IleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLys 918
Db 3103 ATTCTCTTGCTCAATCCGCGGCACCGCTCTTTCTTTCACACAGACATGGGCAAG 3162
Qy 919 AlaLysSerGlyAlaAlaHisLeuPheAlaLeuGluLysLysProAsnIleAspSer 938
Db 3163 GCGAAGAAATCGCGCGCGCGAATTCGACGACTGTTCGACCGAAAGCCACAAATTTGATAAC 3222
Qy 939 ArgSerGlnGluGlyLysProAspThrCysGluGluValPheArgGluVal 958
Db 3223 TGGTCTCGAAGGCGGAGAGCTCGAAGCGTGAAGTGAATTCGAAATTTAGAACGCTG 3282
Qy 959 SerPheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIle 978
Db 3283 CACTTCAGATACCGACCGCCGACAGACAGCTGCTCTCGCGGCTTGGACCTGACCGTG 3342
Qy 979 GluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerVal 998
Db 3343 AAGCTTGCAAAATATGTTGGCTTGTGCGACCCAGCGGTGTGCGCAAGATGACCACTT 3402
Qy 999 GlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAsp 1018
Db 3403 GCATTGCTTGAGCGCTTTTACGATGCGATTCGCGGGTCCATCTTGTGATGGGAAGGAC 3462
Qy 1019 AlaLysLeuLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluPro 1038
Db 3463 ATAAGTAACTAATATCACTCTACCGCAGCTTCTGTCACTGTGTGCGCAGCAGCGCG 3522
Qy 1039 ValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValVal 1058
Db 3523 ACACGTACACGCGGCACCATCAAGAAACATCTTACTTGTATGTGCAAGATGACGTA 3582
Qy 1059 ProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGly 1078
Db 3583 CCGGAAGAATTTTGATTAAGCTTGGCAAGCAGCTAATATCTACGACTTCATCATGTGCG 3642
Qy 1079 LeuProGluLysTyrAsnThrGlnValGlyLysGlyAlaGlnLeuSerGlyGlyGln 1098
Db 3643 CTCCCGGAGGCTTATACAGTTGTTGGCAGCAGGAGGATGTTGTTCTGGCGGCGCAA 3702
Qy 1099 LysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAsp 1118
Db 3703 AAGCAAGCTGTGGCATTGCGCGAGCCCTTCTCGGATCCCAAAATCTTCTCTCGAT 3762
Qy 1119 GluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLys 1138
Db 3763 GAACGAGCTGAGCCCTCGACTCCGAGTCAGAAAGTGTCTCCAGGCGGCTTGGATGCC 3822
Qy 1139 AlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAla 1158
Db 3823 GCTCCCGAGGCGCCACCAATCGCGTTGCACACCGACTCAGCAGCATTCACAAAGCGC 3882
Qy 1159 AspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeu 1178
Db 3883 GACGTTATCTATGTTTTCGACCAAGGCAAGATCGTCGAAAGCGGAACGCAAGCAAGCTG 3942

Qy 1179 LeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerVal 1194
Db 3943 GTCCAGAAAAGCGCGGTACTACGAGCTGGTCAACTTGCAGAGCTTG 3990

RESULT 13

US-08-996-545-3
; Sequence 3, Application US/08996545
; Patent No. 5928898
; GENERAL INFORMATION:
; APPLICANT: Skatrud, Paul L.
; APPLICANT: de Waard, Maarten A.
; APPLICANT: Peery, Robert B.
; APPLICANT: Andrade, Alan C.
; TITLE OF INVENTION: Multiple Drug Resistance Gene atxD of
; TITLE OF INVENTION: Aspergillus nidulans
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: U.S.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/996,545
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Webster, Thomas D.
; REGISTRATION NUMBER: 39,872
; REFERENCE/DOCKET NUMBER: X-11766
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3334
; TELEFAX: 317-276-2763
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4002 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-996-545-3

Alignment Scores:
Pred. No.: 7,23e-247 Length: 4002
Score: 2282.50 Matches: 494
Percent Similarity: 59.70% Conservatives: 232
Best Local Similarity: 40.62% Mismatches: 446
Query Match: 37.97% Indels: 45
DB: 2 Gaps: 11

US-09-873-409-6 (1-1195) x US-08-996-545-3 (1-4002)

Qy 15 LeuProLeuMetProLeuValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu 34
Db 368 GUGCAUUGCGCGCGUGCUUUCAGG-----GGAUUAUGUUAU 409
Qy 35 ValGlnThrAsnThrTyrSerPhePhe-----ArgLeuThrLeuTyrTyrVal 50
Db 410 AUCAAAUUCUGUACGAC-GAGUUCUAUGAUAUCCAAAGACCGUACUGUACUGUA 468
Qy 51 GlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeuTyrIleThr 70
Db 469 UACCUCGGUAUCCGCGGAGUUUGUACUGUACUGUACUGUACUGUACUUAUACC 528

Db 2683 ACAGCCUAGGAGCGGCUAUAUUGCCUGGCAUUGGUGGAAUUGCCUUAUUG 2742
Qy 779 ILeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThr 798
Db 2743 UGUUAUCUGGUGGCGGUGUUCUGGCGGUGGUGGUGGUGGUGGUGGUGGUGG 2802
Qy 799 GlyPheAlaAsnLysAspTyrGlnLeuLysHisAlaGlyLysIleAlaThrGluAla 818
Db 2803 CAGUUCAUACACGCUCAAGGUGUUAUGAGGGAUGUCAAACUUGUUGCGAGGCU 2862
Qy 819 LeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetThr 838
Db 2863 ACAUCGUCUACCGCACAGUUGGUGUUAUACCGGGAAGGUGUUGGGAUUAUAC 2922
Qy 839 GluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleLeuGlySer 858
Db 2923 CAUGCCAGCUUGACGACACAGGACGAGGACAGUUAUCUUCUGGUGGUGGUGG 2982
Qy 859 CysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGly 878
Db 2983 UUAUAUGGUGGUGGCGGACGACUUGUUAUUCUGGUGGUGGUGGUGGUGGUGG 3042
Qy 879 AlaTyrIleuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAla 898
Db 3043 GGGACACUUCUGGUGACACGACGAGUAUGACAUUUUUGGUGUUGUUGUUGG 3102
Qy 899 IleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLys 918
Db 3103 AUUCUUAUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGG 3162
Qy 919 AlaLysSerGlyAlaAlaHisLeuPheAlaLeuGluLysLysProAsnIleAspSer 938
Db 3163 GCGAAGAAUGCGCGCGCGGAAUUCGACGACUGUUGGACGACCAAAUUGAUAC 3222
Qy 939 ArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluVal 958
Db 3223 UGGUGUAGAGGCGGAGAGAGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUG 3282
Qy 959 SerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuIle 978
Db 3283 CACUUCAGAACCGGACCGCGCGGACGACGUGUUGGUGGUGGUGGUGGUGGUGG 3342
Qy 979 GluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerVal 998
Db 3343 AAGCGUGACAAUUAUGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGG 3402
Qy 999 GlnLeuLeuGlnArgLeuTyrAppProValGlnGlnValLeuPheAspGlyValAsp 1018
Db 3403 GCAUUGCUAGGCGCUUUAUGAUGGUGGUGGUGGUGGUGGUGGUGGUGGUGG 3462
Qy 1019 AlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluPro 1038
Db 3463 AUAAGUAAACUAAUAUACUUCUACCGGACGUGUUGGUGGUGGUGGUGGUGG 3522
Qy 1039 ValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValVal 1058
Db 3523 ACACUGUACGAGGCGACCAUACAGGAAACAUUUAUUGGUGGUGGUGGUGGUG 3582
Qy 1059 ProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGly 1078
Db 3583 CCGGAAGAAUUCUUAUUAAGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUG 3642
Qy 1079 LeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGln 1098
Db 3643 CUCCGAGGCGCUUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 3702
Qy 1099 LysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAsp 1118
Db 3703 AAGCAACGUGGCGCAUUGCCGAGCCUUCUUGGGAUCCCAAAUUCUUCUUGAU 3762
Qy 1119 GluAlaThrSerAlaLeuAspAsnAspSerGlnLysValValGlnHisAlaLeuAspLys 1138
Db 3763 GAAGCGAGCGUAGCCUUCGACUCCGAGUCAAAGGUGGUGGUGGUGGUGGUGGUG 3822

Qy 1139 AlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAla 1158
Db 3823 GCUGCCGAGGCGGACCAACCAUAUUGGUGGUGGUGGUGGUGGUGGUGGUG 3882
Qy 1159 AspLeuIleValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeu 1178
Db 3883 GAGUUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 3942
Qy 1179 LeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerVal 1194
Db 3943 GUCCAGAAAAGGCGGUGUACGAGGUGGUGGUGGUGGUGGUGGUGGUGGUG 3990

RESULT 14

US-09-328-320-1
; Sequence 1, Application US/09328320
; Patent No. 6228615
; GENERAL INFORMATION:
; APPLICANT: Skatrud, Paul L.
; APPLICANT: de Waard, Maarten A.
; APPLICANT: Peery, Robert B.
; APPLICANT: Andrade, Alan C.
; TITLE OF INVENTION: Multiple Drug Resistance Gene at/d of
; TITLE OF INVENTION: Aspergillus nidulans
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: U.S.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/328,320
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/996,545
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Webster, Thomas D.
; REGISTRATION NUMBER: 39,872
; REFERENCE/DOCKET NUMBER: X-11766
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3334
; TELEFAX: 317-276-2763
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4002 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..4002
US-09-328-320-1

Alignment Scores:
Pred. No.: 7,23e-247
Score: 2282.50
Percent Similarity: 59.70%
Best Local Similarity: 40.62%
Query Match: 37.97%
DB: 4
Length: 4002
Matches: 494
Conservative: 232
Mismatches: 446
Indels: 45
Gaps: 11

US-09-873-409-6 (1-1195) x US-09-328-320-1 (1-4002)

```
Qy 15 LeuProLeuMetProLeuValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu 34
Db 368 GTGCCATTGCTGCGCGTGCAGCTTTCCAGA-----CGATAAATGTTAT 409
Qy 35 ValGlnThrAsnThrTyrSerPhePhe-----ArgLeuThrLeuTyrTyrVal 50
Db 410 ATCAAAATCTCGTACGAC-GAGTTCTATGATGAATTGACCAAGAACGCTACTGTACTTCGTA 468
Qy 51 GlyIleGlyValAlaLeuIlePheGlyTyrIleGlnIleSerLeuTrpIleThr 70
Db 469 TACCTCGGTATCGGCAGTTGTCACTGTCTATGTTAGTACTGTGGCTTCACTATACC 528
Qy 71 AlaAlaArgGlnThrLysArgIleArgLysGlnPhePheHiserValLeuAlaGlnAsp 90
Db 529 GCGAAGACACGCCAGCAGAGATCCGCGAGTATTACCTTTGAGTCTATCTCGCCAGAAC 588
Qy 91 IleGlyTrpPheAspSerCysAspIleGlyGluLeuAsnThrArgMetThr---AspIle 109
Db 589 ATTGGCTATTTTGAATAACTCGGTGCGCGGGAAGTACCACCGGTATAACGCCGATACA 648
Qy 110 AspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGlnAsnMetSerThr 129
Db 649 AACCTTATCAGGATGGCATTTCCGGAAGGTCGGTCTCACTTGTAGTCCCTCGCGACA 708
Qy 130 PheSerIleGlyLeuAlaValGlyLeuValLysGlyTrpLysLeuThrLeuValThrLeu 149
Db 709 TTCGTGACAGCATTCATTATCGCTACGTCAGTCAATACTGGAAAGTTGGCTCTAATTTGCAGC 768
Qy 150 SerThrSerProLeuIleMetAlaSerAlaAlaCysSerArgMetValIleSerLeu 169
Db 769 TCAACAATTGTGGCCCTCGTTCTCACCTACCGCGGTGTTCTCAGTTTATCATCAAGTAC 828
Qy 170 ThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluValLeuSer 189
Db 829 AGCAAAATCGCTTGACAGCTACGCTGCGGCGGCACTGTTGCGGAAGAGGTTCATCAGC 888
Qy 190 SerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgTyrThrGln 209
Db 889 TCCATCAGAAATGCCACAGGTTTGGCACCACAGACAGACTTCCGAGCAGTATGAGTTC 948
Qy 210 AsnLeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeu 229
Db 949 CACTTAGACGAGCTGAGAAATGGGGAACAAAGAACAGATTGTCATGGTTTCATGATT 1008
Qy 230 GlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSer 249
Db 1009 GCGCCATGTTTGGCCCTTATGTACTCGAACTACGGTCTTGGCTTCTGGATGGGTTCCTCGT 1068
Qy 250 LeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSer 269
Db 1069 TTCTCGGTAGATGGT-----GCAGTCGATGTGGGTGATATCTCAGATTCTCATGGCC 1122
Qy 270 ValIleHiserSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIle 289
Db 1123 ATCTTGATCGATCGTCTCTCTTGGGGAACGTTAGTCCAAATGCTCAAGCATTTTACAAAC 1182
Qy 290 AlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsn 309
Db 1183 GCTGTGCGCGCGCCCAAGATATTTGGAACATCGATCGCCAGTCCCCCATTTAGATCCA 1242
Qy 310 PheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnVal 329
Db 1243 TATTGAAACGAGGGAAGACCTCGACCATTTTGGAGGCCACATTGAGTTACCAATGTC 1302
Qy 330 SerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIle 349
Db 1303 AAGCATATTTACCATCATAGACCCGAGGTCAACCGTCACTGAGGAGTGTCTCTGTCAATG 1362
Qy 350 LysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValVal 369
Db 1363 CCGCGTGGAAACAAACCGCTTTTAGTCCGCCCTCTGGCTCTGGAAAGTACGGTGGTC 1422
```


QY 699 GlyGluLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuThrGlnAsp 718
DB 2503 TCCGAGAGACTTATTCGTCGGCGAGACACTGCCTTTCCGACGATCTCCGCTCAAGAC 2562
QY 719 IleAlaTrpPheAspGluLysGluAsnSerThrGlyLeuThrThrIleLeuAlaIle 738
DB 2563 ATTGCTTTCTTACAGGAAGAATAGACCGCGCTCTGACCTCTTTCCTGTCCACC 2622
QY 739 AspileAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAla 758
DB 2623 GAGACGAGCATCTCCGGTGTAGCGGTGACTCTAGCGACGATCTGTAGCACTCC 2682
QY 759 ThrAsnMetGlyLeuSerValIleIleSerPheIleThrGlyTrpGluMetThrPheLeu 778
DB 2683 ACAGCCCTAGGAGCGGTATCATTTATTCCTCGCATTCGGTGGAAATTTGGCTTAGTT 2742
QY 779 IleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThr 798
DB 2743 TGTATCTCGGTTGCGGTTCTCTGGCATCGGTTCTTACCGATCTATATGCTAGCC 2802
QY 799 GlyPheAlaAsnLysAspLysGlnLeuLysHisAlaGlyLysIleAlaThrGluAla 818
DB 2803 CAGTTTCAATCAGCTCCCAAGCTGTTATGAGGATCTCAAACTTGTTCGAGGCT 2862
QY 819 LeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetThr 838
DB 2863 ACATCGTCTATCCGACAGTTGCGTCAATTAACCCGGGAAAGGATGCTCGGAGATTAC 2922
QY 839 GluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySer 858
DB 2923 CATGCCAGCTTGACGACACAGGACGACAGCTCTAATCTCTGCTGTAGGTCATCCCTG 2982
QY 859 CysThrAlaPheSerHisAlaPheIleThrPheAlaThrAlaAlaGlyPheArgPheGly 878
DB 2983 TTATATGCTGTCGCGAGCACTTGTTTCTTCTGCGTCCGCTGGTTCGTGACGA 3042
QY 879 AlaThrIleuLeuGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAla 898
DB 3043 GGGACACTTCTGTCACACGAGTAGTACATTTTCGCTCTTCTGTTTCTTCGCGAG 3102
QY 899 IleAlaThrGlyAlaMetAlaIleGlyThrLeuValLeuAlaProGluThrSerLys 918
DB 3103 ATTCTCTTGTGTCTCAATCCGCGGCGACCGCTCTTCTTCTTCCACGACATGGGCAAG 3162
QY 919 AlaLysSerGlyAlaAlaHisLeuPheAlaLeuGluLysLysProAsnIleAspSer 938
DB 3163 GCGAAGATCCGCGCGCGAATTCGAGAGTGTTCGCGAAGCCACAAATGTATAAC 3222
QY 939 ArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluVal 958
DB 3223 TGGTCTGAAGAGGCGGAGAGCTCGAAACGCTGGAAGTGAATCGAATTTAGGAACGTG 3282
QY 959 SerPhePheThrProCysArgProAspValPheIleLeuArgGlyLeuSerIle 978
DB 3283 CACTTCAGATACCGACCCCGCCAGAACGCTCTCTGCGCGCTTGGACCTGACCGCTG 3342
QY 979 GluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerVal 998
DB 3343 AAGCTTGGAACAATATGTTCGCTTTCGAGCCAGCGGTTGTGCAAGATGATACCACTT 3402
QY 999 GlnLeuGlnArgLeuThrArgProValGlnGlnValLeuPheAspGlyValAsp 1018
DB 3403 GCATTCCTGAGCGCTTTTACGATGCGATTGCGGGTCCATCTCTTGTATGGAAGGAC 3462
QY 1019 AlaLysGluLeuAsnValGlnThrLeuArgSerGlnIleAlaIleValProGlnGluPro 1038
DB 3463 ATAAGTAACATAATATCAACTCTCTACCGCAGCTTCTGTCTGCTGCTGACCCAGGACCG 3522
QY 1039 ValLeuPheAsnCysSerIleAlaGluAsnIleAlaThrGlyAspAsnSerArgValVal 1058
DB 3523 ACATGTACCGAGGCGCCATCAGAGAAACATCTTACTTGGTATGTCGAAGATGACGTA 3582
QY 1059 ProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGly 1078

DB 3583 CCGAAGAAATCTTGTATTAGGCTTTCAGAGCGCTTAATATCTACGACTTCATCTGTCG 3642
QY 1079 LeuProGluLysThrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGln 1098
DB 3643 CTCGCCGAGGGCTTTATACAGTTGTTGGCAGCAAGGAGGATGTTGTCTGCGGCCAA 3702
QY 1099 LysGlnArgLeuAlaIleAlaArgAlaLeuGlnLeuLysProLysIleLeuLeuAsp 1118
DB 3703 AAGCAACGTTGTGCCATTGCGCGAGCCCTTCTTCGGGATCCCAAAATCTTCTTCTCGAT 3762
QY 1119 GluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLys 1138
DB 3763 GAAGCGACGCTCAGCCCTCAGCTCCGAGTCAGAAAGGTCGTCAGCGCGCTTTGGATGCC 3822
QY 1139 AlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAla 1158
DB 3823 GCTGCCCGAGGCGGACCAACAATCGCGTTGCACCGGACTCAGCACGATTCAAAAGGCG 3882
QY 1159 AspLeuIleValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeu 1178
DB 3883 GAGTTATCTATGTTTTCACCAAGCGCAAGATCGTCGAAAGCGGACGACGACGACTG 3942
QY 1179 LeuArgAsnArgAspIleThrPheLysLeuValAsnAlaGlnSerVal 1194
DB 3943 GTCCAGAAAAGGCGCGGTACTACGAGCTGGTCAACTTCGACAGCTTG 3990
RESULT 15
US-09-328-320-3
; Sequence 3, Application US/09328320
; Patent No. 6228615
; GENERAL INFORMATION:
; APPLICANT: Skatrud, Paul L.
; APPLICANT: de Waard, Maarten A.
; APPLICANT: Peery, Robert B.
; APPLICANT: Andrade, Alan C.
; TITLE OF INVENTION: Multiple Drug Resistance Gene atzD of
; TITLE OF INVENTION: Aspergillus nidulans
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: U.S.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/328,320
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/996,545
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Webster, Thomas D.
; REGISTRATION NUMBER: 39,872
; REFERENCE/DOCKET NUMBER: X-11766
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3334
; TELEFAX: 317-276-2763
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4002 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: mRNA
; HYPOTHETICAL: NO

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 31, 2003, 13:52:37 ; Search time 196.92 Seconds
(without alignments)
2337.888 Million cell updates/sec

Title: US-09-873-409-7
Perfect score: 2739
Sequence: 1 MILGILASLVNGACILPLMLP.....SEKSVAQALKEKTPRYSF 541

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 574371 seqs, 425486471 residues

Total number of hits satisfying chosen parameters: 1148742

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -MODEL=frame+ p2n.model -DEV=xlh
-Q/cgn2_1/USPTO_spool/US09873409/runat_27032003.115424_19290/app_query.fasta_1.7544
-DB=Published Applications NA -QMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09873409 @CGN 1.1 960 @runat_27032003.115424_19290
-NCPU=6 -ICPU=3 -NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREATS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
- 14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	2737	99.9	2021	US-09-873-409-15
2	2697	98.5	3702	US-09-873-409-13
3	2568.5	93.8	1940	US-09-873-409-16
4	2528.5	92.3	3621	US-09-873-409-14

5	1868	68.2	3177	10	US-09-873-409-12	Sequence 12, Appl
6	1566.5	57.2	4189	10	US-09-866-866A-5	Sequence 5, Appl
7	1555.5	56.8	3860	10	US-09-866-866A-1	Sequence 1, Appl
8	1552.5	56.7	4317	9	US-10-044-671-1	Sequence 1, Appl
9	1548	56.5	4254	10	US-09-917-800A-1424	Sequence 1424, Ap
10	1546.5	56.5	3860	10	US-09-866-866A-3	Sequence 3, Appl
11	1546.5	56.5	8630	10	US-09-306-417-1	Sequence 1, Appl
12	1546.5	56.5	8630	10	US-09-306-417-2	Sequence 2, Appl
13	1543	56.3	4425	10	US-09-769-097-3	Sequence 3, Appl
14	1541	56.3	4369	10	US-09-769-097-1	Sequence 1, Appl
15	1539	56.2	4643	9	US-10-072-621-2	Sequence 2, Appl
16	1536	56.1	4788	10	US-09-866-866A-7	Sequence 7, Appl
17	1455.5	53.1	3924	10	US-09-880-107-2299	Sequence 2299, Ap
18	1421.5	51.9	3912	10	US-09-917-800A-1560	Sequence 1560, Ap
19	1220	44.5	5010	10	US-09-917-800A-483	Sequence 483, App
20	1141.5	41.7	1175	10	US-09-873-409-11	Sequence 11, Appl
21	1101.5	40.2	2856	10	US-09-873-409-10	Sequence 10, Appl
22	1028	37.5	1810	10	US-09-749-340-4	Sequence 4, Appl
23	1028	37.5	4175	10	US-09-749-340-3	Sequence 3, Appl
24	993.5	36.3	3861	9	US-09-938-842A-2237	Sequence 2237, Ap
25	942	34.4	3861	9	US-09-938-842A-263	Sequence 263, App
26	914	33.4	4002	9	US-09-758-828-1	Sequence 1, Appl
27	914	33.4	4002	9	US-09-758-828-3	Sequence 3, Appl
28	891.5	32.5	4653	9	US-10-101-388-2	Sequence 2, Appl
29	812.5	29.7	2066	10	US-09-873-409-9	Sequence 9, Appl
30	785.5	28.7	3792	9	US-09-882-694-10	Sequence 10, Appl
31	754	27.5	2681	10	US-09-749-340-8	Sequence 8, Appl
32	754	27.5	8036	9	US-10-101-388-1	Sequence 1, Appl
33	743	27.1	3512	10	US-09-749-340-7	Sequence 7, Appl
34	738	26.9	2698	10	US-09-749-340-5	Sequence 5, Appl
35	688.5	25.1	3999	9	US-09-882-694-9	Sequence 9, Appl
36	657.5	24.0	2298	9	US-10-156-239-18	Sequence 18, Appl
37	657.5	24.0	2298	9	US-09-795-693-18	Sequence 18, Appl
38	657.5	24.0	3408	9	US-10-156-239-16	Sequence 16, Appl
39	657.5	24.0	3408	10	US-09-795-693-16	Sequence 16, Appl
40	657.5	24.0	3512	9	US-10-072-621-1	Sequence 1, Appl
41	596	21.8	2674	10	US-09-917-800A-1500	Sequence 1500, Ap
42	586.5	21.4	1749	10	US-09-815-242-9908	Sequence 9908, Ap
43	582.5	21.3	1749	10	US-09-815-242-6018	Sequence 6018, Ap
44	579.5	21.2	2247	9	US-09-870-759-40	Sequence 40, Appl
45	574.5	21.0	1737	10	US-09-815-242-8403	Sequence 8403, Ap

ALIGNMENTS

RESULT 1
US-09-873-409-15
; Sequence 15, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
; APPLICANT: Frank, Markus
; APPLICANT: Sayegh, Mohamed
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
; TITLE OF INVENTION: Homologue on Chromosome 7p15-21 and Uses Thereof
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873,409
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
; LENGTH: 2021
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: No. US20020037522A1e
; LOCATION: (723)..(723)
; OTHER INFORMATION: n at position 723 represents any nucleotide (A, T, C or G)
US-09-873-409-15

Alignment Scores:
Pred. No.: 3.3e-288 Length: 2021
Score: 2737.00 Matches: 541
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00%
Query Match: 99.93%
DB: 10
Indels: 0
Caps: 0
Mismatches: 0

US-09-873-409-7 (1-541) x US-09-873-409-15 (1-2021)

QY 1 MetileuGlyleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
DB 34 ATGATCCTGGGTACTGGCATCACTGGTCAATGGAGCGCTTCCTTTAAATGCCACTG 93
QY 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
DB 94 GTTTTAGGAGAAATAGGTAGTAACCTTATTAGTGGATGTCTAGTCCAAATTAACACATAC 153
QY 41 SerPhePheArgLeuThrLeuTyrValGlyIleGlyValAlaLeuLeuPheGly 60
DB 154 TCCTTCTTCAGTTGACCTGTATTATGTTGGAATAGTGTGCTGCTTGATTTTGGT 213
QY 61 TyrlleGlnleSerLeuTrpIleleThrAlaAlaArgGlnThrLysArgIleArgLys 80
DB 214 TACATACAGATTTCTTGTGGATTATTAACCTGCAGCAGACCAAGAGGATTCGAAAA 273
QY 81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGly 100
DB 274 CAGTTTTTTTCATTTCAGTTTGGCAGCAGACATCGGCTGGTTTGATAGCTGTGACATCGGT 333
QY 101 GluLeuAsnThrArgMetThrAspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
DB 334 GAACCTTAACACTCGCATGACAGACATTTGACAAATCAGTGTGATTTGGAGATAAGATT 393
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
DB 394 GCTCTGTGTTTCAAAACATGCTACTTTTTCGATTGGCTGGCAGTTGGTTGGTGAAG 453
QY 141 GlyTrpLysleuThrLeuValThrLeuSerThrSerProleuIleMetAlaSerAlaAla 160
DB 454 GGCTGGAAACTCACCTTAGTACTCTATCCAGCTCTCCTCTTAATAGGCTTCAGCGCA 513
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
DB 514 GCATGTTCTAGGATGTCATCTCATTGACCACTAAGGAATTAAGTGCTCTATTCAAGCT 573
QY 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
DB 574 GGGCTGTGCAGAGAAGCTTGTGTCATCAATCCGAACAGTCATAGCTTTAGGGCCAG 633
QY 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
DB 634 GAGAAAGAACTTCAAGGCTCTTCTCTTTTAAATATATACAAAGATATGCTGTTTATTTT 693
QY 221 ProGlnTrpLeuLeuSerCysValLeu**PheValArgTyrThrGlnAsnLeuLysAsp 240
DB 694 CCCAGTGGCTACTAAGTTGGTCTGTTTGTGTTAGGTATACACAGAATCTCAAGAT 753
QY 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
DB 754 GCAAGGATTTGGCAATAAAGAGACTATAGCTTCAAGAGTCTCTCTGGTGTGTAC 813
QY 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
DB 814 TTCTTTATGAATGGAACCTATGGAATGCTTTTGGTATGGAACCTCTCTGATTTCTTAAT 873
QY 281 GlyGluProGlyThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
DB 874 GGAGAACTGGATATACATCGGAGCTGTTCTGCTGTTTCTTTAGTGAATCCATAGC 933
QY 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
DB 934 AGTATTGCAATGGAGCAGCAGTCCCTCACTTGAACCTTCGAATAGCCCGAGGAGCT 993
QY 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
DB 994 GCCTTTTCATATTTTCCAGGTTATTGATAAGAAACCCAGTAGATATACCTTTTCCACAGCT 1053

QY 341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
DB 1054 GGATATAAACCTGAAATCCATAGAAGCACTGTGGAATTTAAAAAATGTTCTTCAATTAT 1113
QY 361 ProSerArgProSerIleLysLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
DB 1114 CCATCAAGACCACTCTATCAAGATTCTGAAAGGTCTGAATCTCAGAAATTAAGCTGGAGAG 1173
QY 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuGln 400
DB 1174 ACAGTCCGCTTGGTGGTCTCAATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAG 1233
QY 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGlyAsnAspIleArgAlaLeu 420
DB 1234 AGCTTATATGATCCGATGATGGCTTTATCATGGTGGATGAGATGCATCAGAGCTTTA 1293
QY 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
DB 1294 AATGTGGCGCATTTATCGAGACCATATTTGGAGTGGTGTAGTCAAGAGCCTGTTTGTTCGGG 1353
QY 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspValThrAspGluGluMetGlu 460
DB 1354 ACCACCATCAGTAACCAATATCAAGTATGACCGAGATGTGACTGATGAAGAGATGGAG 1413
QY 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
DB 1414 AGAGCAGCAAGGAGCAAAATGCTATGATTTTATCATGGAGTTTCTTAATAAATTAAT 1473
QY 481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
DB 1474 ACATTGGTGGGAAAAAGAGCTCAATAGTGGAGGGCAGAAACAGAGGATCGCAATT 1533
QY 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleAspGluAlaThrSerAlaLeu 520
DB 1534 GCTCGTGGCTTAGTTCGAACCCCAAGATTTCTGATTTTAGATGAGGTAGCTGTGCCCTG 1593
QY 521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAspThrProArgTyrSer 540
DB 1594 GATTCAGAAAGCAAGTCAGCTGTTCAAGCTGCACTGGAGAGGATACCCCGAGTATTCA 1653
QY 541 Phe 541
DB 1654 TTT 1656

RESULT 2
US-09-873-409-13
; Sequence 13, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
; APPLICANT: Frank, Markus
; APPLICANT: Sayegh, Mohamed
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
; FILE OF INVENTION: Homologue on Chromosome 7p15-21 and Uses Thereof
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873,409
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 13
; LENGTH: 3702
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: No. US20020037522A1e
; LOCATION: (723)..(723)
; OTHER INFORMATION: n at position 723 represents any nucleotide (A, T, C or G)
US-09-873-409-13

Alignment Scores:
Pred. No.: 1,92e-283 Length: 3702
Score: 2697.00 Matches: 534
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.47% Indels: 0
DB: 10 Gaps: 0

US-09-873-409-7 (1-541) x US-09-873-409-13 (1-3702)

```
QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
DB 34 ATGATCCTGGGTATATCGCATCACTGGTCAATGGAGCCTGCTTCTTTAATGCCACTG 93
QY 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
DB 94 GTTTTAGGAGAAATGATGATTAACCTTATAGTGGATGCTAGTCCAACTAACACATAC 153
QY 41 SerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGly 60
DB 154 TCTTCTTCAGGTTCACCTGATATATGTTGGAATAGGTGTTGCTGCTTGTATTTTGT 213
QY 61 TyrIleGlnIleSerLeuTrpIleThrAlaAlaArgGlnThrLysArgIleArgLys 80
DB 214 TACATACAGATTTCTTGTGGATTATAACTGCAGCAGCAGACCAAGAGGATTCCGAAA 273
QY 81 GlnPhePheHiserValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGly 100
DB 274 CAGTTTTTTTCATTCAGTTTTGGCAGAGGACATCGGCTGGTTGTATAGCTGTGCATCGGT 333
QY 101 GluLeuAsnThrArgMetThrAspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
DB 334 GAACTTAACACTCGCATGACAGACATTGACAAATCAGTGATGTTATCGAGATAAGATT 393
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
DB 394 GCTCTGTTGTTTCAAAACATGCTACTTTTTCGATTGGCTTGGCAGTTGGTTGGTGAAG 453
QY 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
DB 454 GCGTGAACACTCACCTAGTACTCTATCCAGCTCTCTTATATATAGGCTTCAGCGGCA 513
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
DB 514 GCATGTTCTAGGATGGTCATCTCATTCACCATTAAGGATTAAGTGCCTATTCGAAGCT 573
QY 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
DB 574 CGGCGCTGTGGCAGAGAAGTCTGTGTCATCAATCCGAACAGTCATAGCCTTTAGGGCCAG 633
QY 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
DB 634 GAGAAAGAAGCTTCAAAAGGCTTTTCCTTTTAAATATAAACAAGATATGCTTGGTTTATTTT 693
QY 221 ProGlnTrpLeuLeuSerCysValLeu**PheValArgTyrThrGlnAsnLeuLysAsp 240
DB 694 CCCAGTGGCTACTAAGTTGTGTTCTGTTNTTGTAAAGGTATACACAGAATCTCAAGAT 753
QY 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
DB 754 GCAAAAGGATTTGGCATAAAGAGCTATAGCTTCAAAAGTGTCTCTGGTGTGTGTAC 813
QY 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
DB 814 TTCTTTATGAATGGAACTTATGGACTTGCTTTTGTGTGGAACCTTCCTTGATTCCTTAAT 873
QY 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
DB 874 GGGAACCTGGATATACCATCGGACTGTCTTCTGCTGTTTCTTTAGTGTATCCATAGC 933
QY 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
DB 934 AGTTATTGCAATGGAGCAGCTCCCTCACTCTTGTAAACCTTTCGAATAGCCCGAGGAGCT 993
QY 321 AlaPheHisIlePheGlnValIleAspLysProSerIleAspAsnPheSerThrAla 340
DB 994 GCCTTTTCATATTTCCAGGTTATTGTAGTAAAGAAACCCAGTATAGATAACTTTTCCAGCT 1053
```

```
QY 341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
DB 1054 GGATATAAACCTGAATCCATAGAGGAACGTGTGAATTTAAATAATGTTCTTCTCAATTAT 1113
QY 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
DB 1114 CCATCAAGACCATCATCATCAAGATTCTGAAGGTCGAAATCTCAGAAATTAAGTCTGGAG 1173
QY 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
DB 1174 ACAGTCGCTTGGTGGTCTCAATGGCAGTGGAGAGTACGGTAGTCCAGCTTCTGCAG 1233
QY 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
DB 1234 AGGTTATATGATCCGATCATGCTTTTATCATGGTGGATGAGAATCACATCAGAGCTTTA 1293
QY 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
DB 1294 AATGTGGGCATTTATCGAGACCATATTTGGAGTGGTTAGTCAAGAGCCTGTTTGTTCGGG 1353
QY 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
DB 1354 ACCACCATCAGTAACATATCAAGTATGACGAGATGATGTGACTGATGACGAGATGGAG 1413
QY 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
DB 1414 ACATTCGAGCAAGGGAAGCAAAATCGGTATGATTTTATCATGGAGTTCCTATAAATTAAT 1473
QY 481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
DB 1474 ACATTCGAGTGGGGAAGGAGCTCAAAATGAGTGGAGGGCAGAAACAGAGGATCGCAATT 1533
QY 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
DB 1534 GTCGTGCTTGTAGTTGAAACCCCAAGATTCGTATTTAGATGAGGCTACGCTGCGCTG 1593
QY 521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
DB 1594 GATTCAAGAAAGCAAGCTCAGCTGTTCAAGCTGCACCTGGAGAAG 1635
```

RESULT 3

US-09-873-409-16
; Sequence 16, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
; APPLICANT: Frank, Markus
; APPLICANT: Sayegh, Mohamed
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
; FILE OF INVENTION: Homologue on Chromosome 7p15-21 and Uses Thereof
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873,409
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 16
; LENGTH: 1940
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-873-409-16

Alignment Scores:

Pred. No.:	6,98e-270	Length:	1940
Score:	2568.50	Matches:	514
Percent Similarity:	95.01%	Conservative:	0
Best Local Similarity:	95.01%	Mismatches:	0
Query Match:	93.78%	Indels:	27
DB:	10	Gaps:	1

US-09-873-409-7 (1-541) x US-09-873-409-16 (1-1940)

```
QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
DB 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
```


Db 94 GTTTTAGGAGAAATAGTGATGAACCTTTATTAGTGGAGTCTAGTCCAAACTAACATAC 153
 Qy 41 SerPhePheArgLeuThrLeuTyrTyrValGlyValAlaAlaLeuIlePheGly 60
 Db 154 TCTTTCTCAGGTGACCTGTATATATGTTGGAATAGGTGTGCTGCTTGTATTTTGGT 213
 Qy 61 TyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLys 80
 Db 214 TACATACAGATTTCCTGTGTGGATTATAACTGCAGACGACGACAGCAAGAGGATTCGAAA 273
 Qy 81 GlnPhePheHiserValLeuAlaGlnApeIleGlyTrpPheAspSerCysAspIleGly 100
 Db 274 CAGTTTTTTCATTCAGTTTTGGACAGGACATCGGCTGGTTTGTATAGCTGTGACATCGGT 333
 Qy 101 GluLeuAsnThrArgMetThrAspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
 Db 334 GAACTTTAACACTCGCATGCAGACATTCAGAAATCAGTGATGGTATTCGAGATAAGATT 393
 Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
 Db 394 GCTCTGTGTGTTTCAAAACATGCTACTTTTTCGATTGGCCTGGCAGTTGGTTGGTGAAG 453
 Qy 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
 Db 454 GGCCTGAAACTCACCTAGTGACTCTATCCAGCTCTCCTCTTATAATGGCTTCAGCGCA 513
 Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
 Db 514 GCATGTTCTAGGATGGTCATCTCATTCAGCAGTAAGAAATTAAGTCGCTATTCGAAGCT 573
 Qy 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
 Db 574 GGGGCTGTGGCAGAGAGCTTGTGTATCAATCCGAACAGTCATAGCCTTTAGGGCCAG 633
 Qy 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
 Db 634 GAGAAAGAACTTCAA----- 648
 Qy 221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
 Db 649 -----AGGTATACAGAAATCTCAAGAT 672
 Qy 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
 Db 673 GCAAGAGATTTTGGCATAAAGAGCTATAGCTTCAAAAGTGCTCTCTGTGTGTGTAC 732
 Qy 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerIleLeuLeuAsn 280
 Db 733 TTCTTTTATGAATGGAACTTATGGACTTGCTTTTGGTATGGAACTCTCTTGATTTCTTAAT 792
 Qy 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
 Db 793 GGAAACCTGGATATACATCGGAGCTGTTCTTGCTGTTTTCTTAGTGTAAATCCATAGC 852
 Qy 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
 Db 853 AGTTATTGATTCGAGCAGCAGCTCCCTCACTTTGAAACCTTCGCAATAGCCCGAGAGCT 912
 Qy 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
 Db 913 GCCTTTTCATATTTTCCAGGTTATTGTATGAAACCCAGTATAGATAACTTTTCCACAGCT 972
 Qy 341 GlyTyrLysProGluSerIleGlyThrValGluPheLysAsnValSerPheAsnTyr 360
 Db 973 GGATATAAACCTCAATCCATAGAGGAACCTGTGGAATTTTAAATAATGTTCTTTCAATTAT 1032
 Qy 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
 Db 1033 CCATCAAGACCATCTATCAGATTCTCAAGAGTCTGAAATCTCAGATTAGTCTGGAGAG 1092
 Qy 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
 Db 1093 ACAGTGCCTTGGTCTGCTCAATGGCAGTGGGAAGAGTAGTCGGTAGTCCAGCTTCTGCAG 1152

Qy 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
 Db 1153 AGGTTATATGATCCGATGATGGCTTTATCATGGTGGATGAGAAATGACATCAGAGCTTTA 1212
 Qy 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
 Db 1213 AATGTGCGGCATTTATCGAGACCATATTTGGAGTGGTTAGTCAAGAGCCTGTTTTGTTCCGG 1272
 Qy 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
 Db 1273 ACCACCATCAGTAACATATCAATATCAAGTATGACCGAGATGATGTGACTGATGAGAGATGGAG 1332
 Qy 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
 Db 1333 AGAGCAGCAAGGAAAGCAAAATCGTATGATTTTATCATCGAGTTTCTTAATAAATTAAT 1392
 Qy 481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
 Db 1393 ACATTGGTAGGGGAAAGAGAGCTCAAAATGATGGAGGGCAGAAACAGAGGATTCGCAATT 1452
 Qy 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
 Db 1453 GCTCGTCCCTTAGTTTCGAAACCCCAAGATTTCTGATTTTAGATGAGGCTAGCTGCGCTG 1512
 Qy 521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
 Db 1513 GATTTCAGAAAGCAAGTACGCTGTTCAAGCTGCACCTGGAGAAG 1554

RESULT 5

US-09-873-409-12
 ; Sequence 12, Application US/09873409
 ; Patent No. US20020037522A1
 ; GENERAL INFORMATION:

; APPLICANT: Frank, Markus
 ; APPLICANT: Sayegh, Mohamed
 ; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
 ; FILE OF INVENTION: Homologue on Chromosome 7p15-21 and Uses Thereof
 ; FILE REFERENCE: 81994/268611
 ; CURRENT APPLICATION NUMBER: US/09/873,409
 ; CURRENT FILING DATE: 2001-06-05
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: Patent in version 3.0
 ; SEQ ID NO 12
 ; LENGTH: 3177
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: No. US20020037522A1e
 ; LOCATION: (198)..(198)
 ; OTHER INFORMATION: n at position 198 represents any nucleotide (A, T, C or G)
 US-09-873-409-12

Alignment Scores:
 Pred. No.: 2,978-193 Length: 3177
 Score: 1868.00 Matches: 370
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 68.20% Indels: 0
 DB: 10 Gaps: 0

US-09-873-409-7 (1-541) x US-09-873-409-12 (1-3177)

Qy 165 MetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAla 184
 Db 1 ATGGTCTATCTCATTTGACCAAGTAAAGAAATTAAGTGCCTATTCCAAAGCTGGCGCTGTGGCA 60
 Qy 185 GluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeu 204
 Db 61 GAAGAAGTCTTTGTTCATCAATCCGAACAGTCATAGCTTTTAGGGCCCCAGGAGAAAGAACTT 120
 Qy 205 GlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeu 224

Db 121 CAAAGGCTTTCTCTTTTAAATATAACAAGATATGCTTGGTTTATTTTCTCCAGTGCTA 180
Qy LeuSerCysValLeu***PheValatqTyrThrGlnAsnLeuLysAspPhe 244
Db 181 CTAAGTTGTCTGCTGTTNTTGTAAAGGTATACACAGAACTTCAAGAGTGCAGAGGATTTT 240
Qy GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsn 264
Db 241 GGCATAAAGGACTATAGCTTCARAAGTCTCTTGGTCTGCTGTACTTCTTTATGAAT 300
Qy GlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsnGlyGluProGly 284
Db 301 GGAACCTATGGACTTGTCTTTTGGTATGGAACCTCTCTTGATTTCTTAATGAGAACCTGGA 360
Qy TyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIle 304
Db 361 TATACCATCGGGACTGTCTTGTCTGTTTCTTTAGTGAATCCATAGCAGTTATGTGAT 420
Qy GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 324
Db 421 GGAGCAGCAGCTCCCTCACTTTGAAACCTTCGCAATAGCCCGAGGAGCTGCTTTCATATT 480
Qy PheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysPro 344
Db 481 TTCAGGTTATTGATAAGAAACCCAGTATAGATAAATCTTTCCACAGCTGGATATAAACCT 540
Qy GluSerIleGluThrValGluPheLysAsnValSerPheAsnTyrProSerArgPro 364
Db 541 GAATCCATAGAAGAACTGTGGAAATTTAAATAATGTTCTTTCAATATATCCATCAAGACCA 600
Qy SerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu 384
Db 601 TCTATCAAGATTCTGAAGGTCTGAATCTCAGAAATTAAGTCTGGAGAGACAGTCGCCCTTG 660
Qy ValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAsp 404
Db 661 GTCGGTCTCAATGGCAGTGGAGAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGAT 720
Qy ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis 424
Db 721 CCGGATGATGGCTTTATCATGGTGGATGAGAAATGACATCAGAGCTTTAAATGTGCGGCAT 780
Qy TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSer 444
Db 781 TATCGACACCATATGAGTGGTGTAGTCAAGAGCTGTGTTGTCGGACCAACCATCAGT 840
Qy AsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArg 464
Db 841 AACAAATATCAAGTATGACGAGATGATGTGACTGATCAAGAGATGGAGAGCAGCAAGG 900
Qy GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly 484
Db 901 GAAGCAATCGGTATGATTTTATCATCGAGTTTCCCTAATAAATTAATATCATGTTGATGGG 960
Qy GluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu 504
Db 961 GAAAGAGAGCTCAATAGTGGAGGCGCAGAAACAGAGATCGCAATGCTGTCGCTTAA 1020
Qy ValArgAsnProTyrIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 524
Db 1021 GTTCGAAACCCCAAGATCTGATTTTATGATGAGGCTACGCTGCTGCTGGATTCAGAAAGC 1080
Qy 525 LysSerAlaValGlnAlaAlaLeuGluLys 534
Db 1081 AAGTCAGCTGTTCAAGCTGCACTGGAGAAG 1110

RESULT 6

US-09-866-866A-5
; Sequence 5, Application US/09866866A
; Patent No. US2002010224A1
; GENERAL INFORMATION:
; APPLICANT: Sorrentino, Brian
; APPLICANT: Schuetz, John

; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
; FILE REFERENCE: 1340-1-021CIP2
; CURRENT APPLICATION NUMBER: US/09/866, 866A
; CURRENT FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 09/584, 586
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: PCT/US99/11825
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: 60/086, 988
; PRIOR FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 4189
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-866-866A-5

Alignment Scores:
Pred. No.: 3,15e-160 Length: 4189
Score: 1566.50 Matches: 307
Percent Similarity: 72.63% Conservative: 99
Best local Similarity: 54.92% Mismatches: 94
Query Match: 57.19% Indels: 59
DB: 10 Gaps: 6

US-09-873-409-7 (1-541) x US-09-866-866A-5 (1-4189)

Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 148 ATGATTCCTGGAACTCTCGCTGCTATTATCATGGAACATTACTCCCTCTTTGATGCTG 207
Qy 21 ValLeuGlyGluMetSerAspAsnLeuIle----- 30
Db 208 GTCTTTGGAAACATGACAGATAGTTTACAAAGCAGAACCCAGTATTCTGCCAAGCAATT 267
Qy 31 -----SerGlyCysLeuValGlnThrAsnThr----- 39
Db 268 ACTAATCAAAAGTGGACCCCAACAGTACTCTGATCATCAGCAACAGCAGTCTGGAGGAAGAG 327
Qy 40 -----TyrSerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAla 56
Db 328 ATGCGCATATACGCCCTAC-----TATTACACCGGGATTGGTCTGCTGTGTG 372
Qy 57 LeuIlePheGlyTyrIleGlnIleSerLeuTyrIleIleThrAlaAlaArgGlnThrLys 76
Db 373 CTATAGTTGCCCTACATCCAGGTTTCATTTGGTCTGGCAGCTGGAAGACAGATACAC 432
Qy 77 ArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSer 96
Db 433 AAGATTAGGCAGAAAGTTTTCATGCTATAATGAATCAGGAGATAGGCTGGTTTGATGTG 492
Qy 97 CysAspIleGlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGly 115
Db 493 CATGATGTTGGGAGCTCAACACCGGCTCACAGATGATGCTCTCCAAAATTAATACGGA 552
Qy 116 IleGlyAspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAla 135
Db 553 ATTTGGTGACAAAATTTGGGATGTTTTTCAGTCCATCAACCAATTTTAGCCGGTTTTATC 612
Qy 136 ValGlyLeuValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIle 155
Db 613 ATAGGATTTATAAGTGGTGGAGAGCTAACCTTGTCTATTGCTGTCAGCCCTCTTATT 672
Qy 156 MetAlaSerAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSer 175
Db 673 GGATTGTCATCTCTTTTGGGCAAAAGGATTGACTTCACTTACTTAATAGGAACCTCCAG 732
Qy 176 AlaTyrSerLysAlaGlyAlaValAlaGluValLeuSerSerIleArgThrValIle 195
Db 733 GCTTATGCAAAAGCTGGAGCAGTGTCTGAAGAAGTCTTAGCAGCCATCAGAACTGTGATT 792
Qy 196 AlaPheArgAlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyr 215

Db 793 GCCTTTGGAGGCAACAGAGGAAGTGA----- 822
Qy 216 AlaTTPheTyR-PheProGlnTTPLeuLeuSerCysValLeu***PheValArgTyrThr 235
Db 823 -----AGGTACAT 831
Qy 236 GlnAsnLeuLysAspAlaLysAspPheGlyLeuLeuArgThrIleAlaSerLysValSer 255
Db 832 AAAAATTTAGAGAGCTAAATATGTTGGCATAAAGAAAGCTATCACAGCCAGCATTTGC 891
Qy 256 LeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrGlyThr 275
Db 892 ATAGGCATTGCCCTGTTGGTCTATGCATCATATGCCTGGCATTTCTGGTATGGGACA 951
Qy 276 SerLeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePhe 295
Db 952 TCTTGGTCTCTCAATGAA-----TATTCTATTGGAGAGTGTCTTACTGCTCTTC 1005
Qy 296 SerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAla 315
Db 1006 TCTATTGTTGGGACTTTTAGTATTGGACACTTGGCCCAACATAGAGCTTTGCA 1065
Qy 316 IleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAsp 335
Db 1066 AAGCGACAGGGCGACCTTTGAAATCTTCAAGATAATTGATAACGAGCCACGATTGAC 1125
Qy 336 AsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsn 355
Db 1126 AGCTTCTCAACAGGGCTACAAACAGACAGATATATGGGAACCTTAGAGTTTAAAT 1185
Qy 356 ValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArg 375
Db 1186 GTTCACTTCAACTACCCATCGACAGAGGAGTTTCAGATCTTGAAGGCTTCAATCTGAAG 1245
Qy 376 IleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrVal 395
Db 1246 GTCAAGAGCGGACAGACGGTGGCTTGGTGGCAACAGTGGCTGTGGAAAAAGCAAACT 1305
Qy 396 ValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsn 415
Db 1306 GTCCACCTGATGCAGAGGCTCTACGACCCCTCGAGGGCGGTGTCTAGTATCGAGGACAA 1365
Qy 416 AspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGlu 435
Db 1366 GACATCAGAACCATCAATGTGAGGTATCTCAGGAGAGATCATTTGGTGTGAGTCAGGAA 1425
Qy 436 ProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThr 455
Db 1426 CCGTGTCTGTTGCCACCCAGCATCGCCGAGAACATTCGCTATGCGCGAGAGATGTCAACC 1485
Qy 456 AspGluGluMetGluArgAlaAaArgGluAlaAsnAlaTyrAspPheIleMetGluPhe 475
Db 1486 ATGGATGAGATTCAGAAAGCTGTCAAGGAAGCCAAATGCCCTATCACTTCATCAAGAACTG 1545
Qy 476 ProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLys 495
Db 1546 CCCACCAATTCACACCTGTTGGTGTGAGAGAGGGCGGCGAGTGTGGGGGACAGAAA 1605
Qy 496 GlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuLeuLeuAspGlu 515
Db 1606 CAGAGAATGCCATTCGCCGGCCCTGTCGCGCAATCCCAAGATCCCTTTGTTGGACGAG 1665
Qy 516 AlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaLeuGluLys 534
Db 1666 GCCACCTCAGCCCTGGATACAGAAAGTGAAGCTGTGTGTGCGAGCCGACCTGGATAAG 1722

RESULT 7

US-09-866-866A-1
; Sequence 1, Application US/09866866A
; Patent No. US20020102244A1
; GENERAL INFORMATION:
; APPLICANT: Sorrentino, Brian

; APPLICANT: Schuetz, John
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
; FILE REFERENCE: 1340-1-021CIP2
; CURRENT APPLICATION NUMBER: US/09/866,866A
; CURRENT FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 09/584,586
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: PCT/US99/11825
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: 60/086,988
; PRIOR FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 3860
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-866-866A-1

Alignment Scores:

Pred. No.: 4,38e-159 Length: 3860
Score: 1555.50 Matches: 306
Percent Similarity: 73.29% Conservative: 100
Best Local Similarity: 55.23% Mismatches: 99
Query Match: 56.79% Indels: 49
DB: 10 Gaps: 6

US-09-873-409-7 (1-541) x US-09-866-866A-1 (1-3860)

Qy 1 MetIleLeuGlyIleleuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 151 ATGGTGGTGGGAACCTTTGGCTGCCATCATCCATGGGGCTGCCTCTCTCATGATGCTG 210
Qy 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
Db 211 GTGTTGGAGAAATGACAGATATCTTTGCCAAATGCAGGAAATTTAGAACATCTGATGTC 270
Qy 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe 43
Db 271 AACATCACTAATAGAAGTGATATCAATGATACAGGGTCTTTCATGAATCTGGAGGAAGAC 330
Qy 44 -----ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db 331 ATGACACAGATATGCCCTATTATTATACAGTGGAAATTTGGTGGGGTGTCTGCTGCTTAC 390
Qy 62 IleGlnIleSerLeuTyrIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db 391 ATTCAGGTTTTCATTTTGGTGGCTGGCAGCTGGAAGCAAAATACACAAAATTAGAAAACAG 450
Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101
Db 451 TTTTTCATGCTATAATGCGACAGAGATAGCTGTTTGTATGTGCACGATGTTGGGGAG 510
Qy 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db 511 CTTAACACCCGAGTACAGATGATGTCTTAAGATTAAATGAAGGTATTCGTGCACAAAAT 570
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db 571 GGATGTTCTTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGATTACACGT 630
Qy 141 GlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db 631 GGTGGGAAGCTAACCCCTGCTGATTTTGGCCATCAGTCTCTGTTCTTCTGCAGCTGCT 690
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerLysAla 180
Db 691 GTCTGGGCAAGATATCTATCTTCAATTTACTGATAAAGAACTCTTCTAGCGTATGCAAAAGCT 750
Qy 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db 751 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCAATTTGGAGGACAA 810

201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
Db 811 AAGAAAGAACTTGAA----- 825
Qy 221 ProGlnTrpLeuLeuSerCysValLeu**PheValArgTyrThrGlnAsnLeuLysAsp 240
Db 826 -----AGGTACACAAATAATTAGAGAA 849
Qy 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
Db 850 GCTAAAAGAAATGGGATAAGAAAGATTACAGCCAAATTTCTATAGTGTCTGCTTTC 909
Qy 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrPyrGlyThrSerLeuIleLeuAsn 280
Db 910 CTGCTGATCTATGCATTTATGCTCTGGCTCTGGTATGGACCACTTGTTCCTCA 969
Qy 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
Db 970 GGGGAA-----TATTCTATTGGACAAGTACTCACTGTATTCTTTCTGTATTAAATGGG 1023
Qy 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
Db 1024 GCITTTAGTGTGGACAGGATCTCCAGGATTCGAAGCATTTGCAATGCAAGAGGACGA 1083
Qy 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
Db 1084 GCITATGAATCTTCAAGATAATTTGATAAAGCAAGTATTGACAGCTATTGGAAGAGT 1143
Qy 341 GlyTyrLysProGluSerIleGlyThrValGluPheLysAsnValSerPheAsnTyr 360
Db 1144 GGGCACAACACAGATAATATTAAGGGAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 1203
Qy 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
Db 1204 CCATCTCGAAAAGAAATTAAGATCTTGAAGGCTGAACTGGAAGTGGAGAGTGGGCG 1263
Qy 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuGln 400
Db 1264 ACGTGGCCCTGGTGGAAACAGTGGCTGGTGGGAAGACACACAGTCCAGCTGATGACG 1323
Qy 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
Db 1324 AGGCTCTATGACCCACAGAGGGGATGCTCAGTGTGATGGACAGGATATTAGGACCATA 1383
Qy 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
Db 1384 AATGTAAGGTTTCTACGGGAATCAITTTGGTGTGTGATGAGTCAAGAACCTGTATTGTTGCC 1443
Qy 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspValThrAspGluGluMetGlu 460
Db 1444 ACCAGATAGCTGAAACATTCGCTATGGCCGTGAAATGTCAACATGGATGAGATTGAG 1503
Qy 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
Db 1504 AAAGCTCAAGGAAGCAATGCTATGACTTTATCATGAACTGCCTCATATAAATTGAC 1563
Qy 481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlnLysGlnArgIleAlaIle 500
Db 1564 ACCCTGTTGGAGAGAGAGGGGCCATTTAGTGTGGGAGAGAGAGATGCGCCATT 1623
Qy 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
Db 1624 GCAGTGGCCCTGGTTCGCAACCCCAAGATCTCTCTGCTGGATGAGGACGCTGAGCCTTG 1683
Qy 521 AspSerGluSerLysSerAlaValGlnAlaLeuGluLys 534
Db 1684 GACACAGAAACGAAGCAGTGGTTTCAGTGGCTCTCGGATAAG 1725

RESULT 8

US-10-044-671-1
; Sequence 1, Application US/10044671
; Patent No. US20020177147A1
; GENERAL INFORMATION:

APPLICANT: Washington State University Research Foundation
APPLICANT: Mealey, Katrina
APPLICANT: Bentjen, Steven
TITLE OF INVENTION: MDR1 VARIANTS AND METHODS FOR THEIR USE
FILE REFERENCE: 4630-61733
CURRENT APPLICATION NUMBER: US/10/044,671
CURRENT FILING DATE: 2002-01-10
PRIOR APPLICATION NUMBER: US 60/261,578
PRIOR FILING DATE: 2001-01-12
PRIOR APPLICATION NUMBER: US 60/314,829
PRIOR FILING DATE: 2001-08-24
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 4317
TYPE: DNA
ORGANISM: Canis familiaris
FEATURE:
NAME/KEY: CDS
LOCATION: (70)..(3912)
OTHER INFORMATION:
US-10-044-671-1
Alignment Scores:
Pred. No.: 1,le-158 Length: 4317
Score: 1552.50 Matches: 302
Percent Similarity: 71.94% Conservative: 98
Best Local Similarity: 54.32% Mismatches: 105
Query Match: 56.68% Indels: 51
DB: Gaps: 5
US-09-873-409-7 (1-541) x US-10-044-671-1 (1-4317)

Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyValaCysLeuProLeuMetProLeu 20
Db 217 ATGTTGGTGGGACAATGGCTGCCATCTCCATGGAGCTGCATCCCTCTCATGATGCTG 276
Qy 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu----- 34
Db 277 GTTTTGGAAACATGACAGATAGCTTTGCCAATGCAGGAATTTCAAGAAACAAACTTTT 336
Qy 35 -----ValGlnThrAsnThrTyrSerPhePheArg----- 44
Db 337 CCAGTTATAATTAATGAAGTATTACGAACAATACACAACATTTTCATCAACCATCTGGAG 396
Qy 45 -----LeuThrLeuTyrValGlyIleGlyValaAlaLeuIlePhe 59
Db 397 GAGGAATGACACGATGCTGCTATTATTACAGTGGATCGGTGCTGGTGGTGGCT 456
Qy 60 GlyTyrIleGlnIleSerLeuTyrPheIleThrAlaAlaArgGlnThrLysArgIleArg 79
Db 457 GCTTACATCAGGTTTCATCTGCTGCTGGCAGCAGGAGACAGATACTCAAAATTAGA 516
Qy 80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99
Db 517 AAACAATTTTTCATGCTATCATGCGACAGGAGATGGCTGGTTTGCAGTGCATGACGTT 576
Qy 100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
Db 577 GGGGAGCTTAACACCGGCTCACAGACGATGCTCTCAAAATCAATGAAGAAATGGCGAC 636
Qy 119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
Db 637 AAGTTGGATGTTCTTCAATCAATAGCAACATTTTTCACCGGTTTATAGTGGGTTT 696
Qy 139 valLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
Db 697 ACACCTGGTGGAAAGCTAACCCCTGTGATTTGGCCATCAGCCCTGTCTTGGACATTCA 756
Qy 159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
Db 757 GCCGCCATCGGCAAGATACCTATTCTTATTACTGATTAAGAACTCTTGGCGCTATGCA 816

QY 179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198
DB 817 AARGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGAACTGTGATGCTTGGGA 876
QY 199 AlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPhe 218
DB 877 GGACAAAGAAAGAACTTGA----- 897
QY 219 TyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeu 238
DB 898 -----AGGTACAAACAAATTTA 915
QY 239 LysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAla 258
DB 916 GAAGAAGCTAAAGAAATGGGATAAAGAAAGTATCAGCGCAACATTTCTATTGGTGCC 975
QY 259 ValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIle 278
DB 976 GCTTTCTTATTGATCATGATATGCTCTGGCTTTCTGGTATGGACCTCTTGGTC 1035
QY 279 LeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIle 298
DB 1036 CTCTCAGTGAA-----TATACTATTGGACAGTACTCACTGCTCTCTTTCTGTATT 1089
QY 299 HisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArg 318
DB 1090 ATTGGGCTTTTAGTATTGGACAGCATCCCAAGCATTTGAAGCATTTGCAACGCAAGA 1149
QY 319 GlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSer 338
DB 1150 GGAGCAGCTTATGAATCTTCAAGATAATTGCAATAAACCAAGCATTTGACACTATTGC 1209
QY 339 ThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPhe 358
DB 1210 AAGAGTGGACATAAACAGATATATTAGGGAATTTGGAATTCAAATGTTCACTTC 1269
QY 359 AsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSer 378
DB 1270 AGTTACCTCTCGAAAAGAAAGTTAAGATCTTAAGGGTCTCAACCTGAAGGTTACAGT 1329
QY 379 GlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeu 398
DB 1330 GGGCACACATGCGCTGTGTGGAAACAGTGGCTCGGGGAAGACACGCCGTGCAGCTG 1389
QY 399 LeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArg 418
DB 1390 ATGCAGAGGCTCTATGACCCACAGATGGCATGCTGTGATTGATGACAGGACATTAGG 1449
QY 419 AlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeu 438
DB 1450 ACCATAAATGTAAGGCATCTTCGGGAAATTTACTGTTGTGTGATCAGGAGCCTGTGTTG 1509
QY 439 PheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGlu 458
DB 1510 TTTGCCACACGATAGCTGAAAAACATTCGCTATGGCGCGCAAAATGTCACCATGGATGAG 1569
QY 459 MetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLys 478
DB 1570 ATTGAGAAAGCTGTTTAGGAAGGCAATGCTCATGATTTTATCATGAAGATACCTAATAAA 1629
QY 479 PheAsnThrLeuValGlyGlyLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIle 498
DB 1630 TTTGACACTCTGTTGGAGAGAGAGGGCGCGCTGAGTGGTGGACAGAAACAGAGAAATC 1689
QY 499 AlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSer 518
DB 1690 GCCATTGCTCGGCGCTGTTGTCGCAACCCCAAGATTTCTTCTGCTGGATGAGGCAACGTC 1749
QY 519 AlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
DB 1750 GCTCTGGACACTGAAAGTGAAGCAGATGGTTTCAGGTGGCCCTGGATAG 1797

RESULT 9

US-09-917-800A-1424
; Sequence 1424, Application US/09917800A
; Patent No. US20020119462A1
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Castie, Arthur
; APPLICANT: Elashoff, Michael
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5038-US
; CURRENT APPLICATION NUMBER: US/09/917,800A
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/222,040
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: US 60/222,880
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: US 60/290,029
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/295,798
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/297,457
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,884
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,459
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 1740
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1424
; LENGTH: 4254

TYPE: DNA
ORGANISM: Rattus norvegicus
FEATURE:
OTHER INFORMATION: Genbank Accession No. US20020119462A1 M81855
US-09-917-800A-1424

Alignment Scores:
Pred. No.: 3,34e-158 Length: 4254
Score: 1548.00 Matches: 307
Percent Similarity: 72.25% Conservative: 94
Best Local Similarity: 55.32% Mismatches: 104
Query Match: 56.52% Indels: 50
DB: 10 Gaps: 5

US-09-873-409-7 (1-541) x US-09-917-800A-1424 (1-4254)

QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
DB 247 ATGGCTCTCGGAACCTCGCTCGCTATCATCCACGGAACCTGCTTCCCTCTCGATGCTG 306
QY 21 ValLeuGlyGluMetSerAspAsnLeuIle----- 30
DB 307 GTGTTCCGATACATGACAGATAGTTTACCCNAGCAGAGACCCGCAATTCGACCGAGCG 366
QY 31 -----SerGlyCysLeuValGlnThrAsnThrTyrSerPhePheArgLeu----- 45
DB 367 ATTTACTAATCAAAAGTGAATCAACAGTACACATACCGTCAGCGCACAGAGTCTGGAGGAG 426
QY 46 -----ThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGly 60
DB 427 GACATGCCCATGTATGCCCTACTATTATACGGGCATTTGGTCCCGGTGTGCTCATCTGCC 486
QY 61 TyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLys 80
DB 487 TACATCCAGGTTTCACTTTGGTCCCTCGGAGTGGGAGACAAATACACAAGATTAGGCAG 546
QY 81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGly 100

QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101
 Db 451 TTTTTCATGCTATAATGCACAGGAGATAGGCTGTTGATGTCACCATGTTGGGAG 510
 QY 102 LeuAsnThrArgMetThr--AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
 Db 511 CTTAACACCCGACTTACAGATGATGCTCTAAGATTAAATGAAGTTATTTGGTGACAAAATT 570
 QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
 Db 571 GGAATGCTTTTCAGTCATGCGCAACATTTTTCACCTGGGTTTATAGTAGGATTTACACGT 630
 QY 141 GlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
 Db 631 GGTGTGAAGCTAACCTCTGTGATTTTGGCAGCAATTTAGAACTGTGATTCGATTTGGAGCAAA 690
 QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
 Db 691 GTCGTGGCAAGATACCTATCTTCATTACTGATTAAGAACTCTTAGCGTATGCAAAAGCT 750
 QY 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
 Db 751 GGAAGCAGTAGCTGAAGAGGCTCTGGCAGCAATTTAGAACTGTGATTCGATTTGGAGCAAA 810
 QY 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTyrPheTyrPhe 220
 Db 811 AGAAAGAACTTGAA----- 825
 QY 221 ProGlnTyrLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
 Db 826 -----AGGTACAAACAAATTTAGAAAGAA 849
 QY 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
 Db 850 GCTPAAAGAAATTCGGATTAAGAAAGCTATTACAGCAATTTCTATAGTGCTGCTTTC 909
 QY 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrGlyThrSerLeuLeuLeuAsn 280
 Db 910 CTGCTGATCTATGATCTTATGCTCTGGCCTCTGGTATGGGACCACCTTTGGTCTCTCA 969
 QY 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
 Db 970 GGGGAA-----TATTCTATTTGGAAGAAGTACTCACTGTATTTCTTCTGTTATTAATTTGGG 1023
 QY 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
 Db 1024 GCCTTTAGTGTGGACAGGATCTCCAGCATTTGCAATTTGCAATGCAAGAGAGCA 1083
 QY 321 AlaPheHisIlePheGlnValIleAspLysPheProSerIleAspAsnPheSerThrAla 340
 Db 1084 GCCTTATGAAATCTTCAAGATAATTGATAATAGCCCAAGTATTGACAGCTATTGCAAGAGT 1143
 QY 341 GlyTyrLysProGluSerIleGlyGlyThrValGluPheLysAsnValSerPheAsnTyr 360
 Db 1144 GGGCACAACACAGATAATATTAAGGGGAAATTTGGAATTTGCAATTTGCAATTTGCAATTTAC 1203
 QY 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
 Db 1204 CCATCTCGAAAGAAAGATTAGATCTTGAAGGGCTGAACTGAAGGTGCAAGATGGGAG 1263
 QY 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
 Db 1264 ACGGTGGCCCTGTTGGAAACAGTGGCTGTGGGAAGACACACAGTCCAGCTGATGTCAG 1323
 QY 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
 Db 1324 AGGCTCTATGACCCCAAGAGGGGATGGTCACTGCTGATGACAGGAGTATAGGACCATTA 1383
 QY 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
 Db 1384 AATGTAAGGTTTCTACGGGAAATCATTTGGTGTGAGTCAGGAACCTGTATTGTTGCC 1443
 QY 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460

Db 1444 ACCCAGATAGCTGAAACATTCGCTATGCGCGTGAATAATGTCCACCATGATGAGATTGAG 1503
 QY 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
 Db 1504 AAAGCTGTCTAGAGACCCATGCTTATGACTTATCATGAAACTGCCTCATTAATTTGAC 1563
 QY 481 ThrLeuValGlyGlyLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
 Db 1564 ACCCTGGTTGGAGAGAGAGGGGCCAGTTGAGTGTGGGACAGAGGATCGCCATT 1623
 QY 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
 Db 1624 GCACGTGCTGCTGGTTCGCAACCCCAAGATCCTCTGCTGGATGAGGCCACGTCAGCCTTG 1683
 QY 521 AspSerGluSerLysSerAlaValGlnAlaLeuGluLys 534
 Db 1684 GACACAGAGGCGAAGCAGTGGTTCAGGTGGCTCTGGATAAG 1725

RESULT 11
 US-09-306-417-1
 ; Sequence 1, Application US/09306417
 ; Patent No. US20020103144A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Heinrich-Pette-Institut
 ; TITLE OF INVENTION: Retroviral Gene Transfer Vectors
 ; FILE REFERENCE: P50491
 ; CURRENT APPLICATION NUMBER: US/09/306,417
 ; CURRENT FILING DATE: 1999-05-06
 ; EARLIER APPLICATION NUMBER: DE 198 22 115
 ; EARLIER FILING DATE: 1998-05-08
 ; NUMBER OF SEQ ID NOS: 17
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 1
 ; LENGTH: 8630
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: proviral
 ; OTHER INFORMATION: plasmid DNA
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)..(160)
 ; OTHER INFORMATION: plasmid backbone (pUC)
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (161)..(677)
 ; OTHER INFORMATION: 5'-LTR
 ; FEATURE:
 ; NAME/KEY: 5'UTR
 ; LOCATION: (532)..(1219)
 ; FEATURE:
 ; NAME/KEY: mat_peptide
 ; LOCATION: (1220)..(5062)
 ; OTHER INFORMATION: m4 mdr-1 cDNA
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (5215)..(5774)
 ; OTHER INFORMATION: 3'-LTR
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (5775)..(8630)
 ; OTHER INFORMATION: plasmid backbone (pUC)
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)..(8630)
 ; OTHER INFORMATION: retroviral expression vector SFbeta71m4
 US-09-306-417-1

Alignment Scores:
 Pred. No.: 1,466-157 Length: 8630
 Score: 1546.50 Matches: 305
 Percent Similarity: 73.10% Conservative: 100

Best Local Similarity: 55.05% Mismatches: 100
Query Match: 56.46% Indels: 49
DB: 10 Gaps: 6

US-09-873-409-7 (1-541) x US-09-306-417-1 (1-8630)

QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
DB 1370 ATGGTGGTGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCTCTCATGATGCTG 1429
QY 21 ValLeuGlyGluMetSerAsp-----AnLeuIleSer 31
DB 1430 GTGTTTGGAGAAATACAGATATCTTTGCAAAATGCAGAAATTTAGAGATCTGATGTCA 1489
QY 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
DB 1490 AACATCTAATAGAGATGATATCAATGATACAGGGTTCTTCATGATCTGGAGGAGAC 1549
QY 44 -----ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
DB 1550 ATGACACAGATATGCTTATTATACAGTGGAAATGGTCTGGGGTGGTGGTCTGCTTAC 1609
QY 62 IleGlnIleSerLeuThrIleLeuThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
DB 1610 ATTCAGGTTTCATTTTGGTGGCAGGATAGGCTGGTGGTGGTGGTGGTGGTGGTGGTGG 1669
QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheSerCysAspIleGlyGlu 101
DB 1670 TTTTTCATGCTATATGCGACAGGATAGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGGAG 1729
QY 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
DB 1730 CTTAACACCCGACTTACAGATGATGCTCTTAAGTATTAATGAAGTATTTGGTGGACAAAT 1789
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
DB 1790 GGAATGTTCTTTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGGATTTACACGT 1849
QY 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
DB 1850 GGTTGGAAAGCTAACCCCTGTGATTTTGGCATCATGCTGCTGTCTTGGACATGTCAGCTGCT 1909
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
DB 1910 GTCTGGCGCAAGATACTATCTTCACTTACTCATTAAGAAGCTCTTAGCGTATGCCAAGCT 1969
QY 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
DB 1970 GGAGCAGTAGCTGAAGAGGCTTTGGCAGCAATTAGAAGCTGTGATTCGATTTGGAGGACAA 2029
QY 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
DB 2030 AAGAAAGAACTTGAA----- 2044
QY 221 ProGlnTrpLeuLeuSerCysValLeu**PheValArgTyrThrGlnAsnLeuLysAsp 240
DB 2045 -----AGGTACACAAAAATTTAGAGAA 2068
QY 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
DB 2069 GCTAAAAGAAATGGGATAAAGAAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTC 2128
QY 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
DB 2129 CTGCTGATCTATGATCTTATGCTCTGGCTTCTGGTATGGGACCACTTGGTCTCTCTCA 2188
QY 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
DB 2189 GGGGAA-----TATTCTATGGACAGACTACTGATCTTTCTCTGATTAATTTGGG 2242
QY 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
DB 2243 GCTTTTAGTGTGGACAGGATCTCCAGCAATTTGAAGCAATTTGCAAAATGCAAGAGGACGA 2302

QY 321 AlaPheHisIlePheGlnValIleAspTyrLysProSerIleAspAsnPheSerThrAla 340
DB 2303 GCTTATGAAATCTTCAAGATAATTGATAATAAGCCAAAGTATTGACAGCTATTCCGAAGAGT 2362
QY 341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
DB 2363 GGGCACAACCCAGATAATTAAGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 2422
QY 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
DB 2423 CCATCTCGAAAAGAGTTAAGATCTTGAAGGGCTGAACTGAAGTGCAGAGTGGGCGAG 2482
QY 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
DB 2483 ACGTGGCCCTGGTTGGAAACAGTGGCTGTGGGAAGAGCACACAGCTCCAGCTGTATGCAG 2542
QY 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
DB 2543 AGGCTCTATGACCCACAGAGGGGATGGTCAAGTGTGGTGGACAGGATATTAGGACCAT 2602
QY 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
DB 2603 AATGTAAGGTTTCTACGGGAAATCATTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2662
QY 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspValThrAspGluGluMetGlu 460
DB 2663 ACCACGATAGCTGAAACATTCGCTATGCGCGTGAAGTGTCCACCATGGATGAGATTGAG 2722
QY 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
DB 2723 AAGCTGTCAAGGAAGCAATGCTATGATCTTATCATGAACTGCTCATAAATTTGAC 2782
QY 481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
DB 2783 ACCCTGGTGGAGAGAGAGGGGCCAGTGTGAGTGGTGGGAGAGAGAGAGTGGCCATT 2842
QY 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
DB 2843 GCACGTGGCTGGTGGTGGCAACCCCAAGATCCCTCTGCTGGATGAGGCCACGTCAGCCTTG 2902
QY 521 AspSerGluSerLysSerAlaValGlnAlaLeuGluLys 534
DB 2903 GACACAGAAAGCAAGCAGTGGTTCAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTAA 2944

RESULT 12
US-09-306-417-2
; Sequence 2, Application US/09306417
; Patent No. US20020103144A1
; GENERAL INFORMATION:
; APPLICANT: Heinrich-Pette-Institut
; TITLE OF INVENTION: Retroviral Gene Transfer Vectors
; FILE REFERENCE: P50491
; CURRENT APPLICATION NUMBER: US/09/306,417
; CURRENT FILING DATE: 1999-05-06
; EARLIER APPLICATION NUMBER: DE 198 22 115
; EARLIER FILING DATE: 1998-05-08
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 8630
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: proviral
; OTHER INFORMATION: plasmid DNA
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(8630)
; OTHER INFORMATION: retroviral expression vector SFbeta91msA1
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(160)

OTHER INFORMATION: plasmid backbone (pUC)

FEATURE:

NAME/KEY: misc feature

LOCATION: (161)..(677)

OTHER INFORMATION: 5'-LTR

FEATURE:

NAME/KEY: 5'UTR

LOCATION: (532)..(1219)

FEATURE:

NAME/KEY: mat_peptide

LOCATION: (1220)..(5062)

OTHER INFORMATION: msal mdrl cdna

FEATURE:

NAME/KEY: misc feature

LOCATION: (5215)..(5774)

OTHER INFORMATION: 3'-LTR

FEATURE:

NAME/KEY: misc feature

LOCATION: (5775)..(8630)

OTHER INFORMATION: plasmid backbone (pUC)

US-09-306-417-2

Alignment Scores:

Pred. No.: 1,468-157 Length: 8630
Score: 1546.50 Matches: 305
Percent Similarity: 73.10% Conservative: 100
Best Local Similarity: 55.05% Mismatches: 100
Query Match: 56.46% Indels: 49
DB: 6 Gaps: 6

US-09-873-409-7 (1-541) x US-09-306-417-2 (1-8630)

QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
DB 1370 ATGGTGGGGAACCTTGGTGCATCATCCATGGGGCTGGACTTCTCTCATGATGCTG 1429
QY 21 ValLeuGlyGluMetSerAsp-----AenLeuIleSer 31
DB 1430 GTGTTGGGAAATGCAGATATCTTTGCAAAATGCAGGAAATTTAGAAATCTGATGTC 1489
QY 32 GlyCysLeuValGlnThr-----AenThrTyrSerPhePhe----- 43
DB 1490 AACATCACTAATAGAGTATATCAATGATACAGGGTCTTCTATGATCTGGAGGAGAC 1549
QY 44 -----ArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
DB 1550 ATGACAGGATGCTTATTAATCAGTGAATTTGGTCTGGGTGCTGGTGTCTGCTTAC 1609
QY 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleAryLysGln 81
DB 1610 ATTCAGGTTTCATTTGGTCTGGCAGCTGGAGACAAATACAAAATTAGAAAACAG 1669
QY 82 PhePheHisSerValLeuAlaGlnApsIleGlyTrpPheAspSerCysAspIleGlyGlu 101
DB 1670 TTTTTCATGCTATAATGCACAGGAGATAGGCTGGTTTATGTCACGATGTTGGGAG 1729
QY 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
DB 1730 CTTAACACCCGATACAGATGATGCTCTAAGATTAATGAAGTATTTGGTGCAGAAAT 1789
QY 121 AlaLeuLeuPheGlnAenMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
DB 1790 GGAATGTTCTTCAGTCAATGGCAACATTTTTCATCTGGTTTATAGTATTTACACGT 1849
QY 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
DB 1850 GGTGGGAAGTAAACCTTGTGATTTTGGCCATCAGTCTCTTCTTGGCATGTCAGCTGCT 1909
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGlnLeuSerAlaTyrSerLysAla 180
DB 1910 GTCTGGGCAAGACTACTATCTTACTTACTTAAAGAACTCTTAGCGTATGCAAAAGCT 1969
QY 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200

RESULT 13

US-09-769-097-3

DB 1970 GGAGCAGTAGCTGAAGAGGCTTTGGCAGCAATTAGAACTGTGATTCATTTGGAGACAA 2029
QY 201 GluLysGluLeuGlnArgSerPheLeuLeuAenIleThrArgTyrAlaTrpPheTyrPhe 220
DB 2030 AGAAGAAGACTTGAA----- 2044
QY 221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAenLeuLysAsp 240
DB 2045 -----AGGTACAACAAAAAATTTAGAGAA 2068
QY 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
DB 2069 GCTAAAAGAAATTTGGGATAAAGAAAGCTATTACAGCAATATTCTATAGTGTGCTTTC 2128
QY 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrGlyThrSerLeuIleLeuAsn 280
DB 2129 CTGCTGATCTATGATCTTATGCTCTGGGCTTCTGGTATGGGACACCTTGGTCTCTCA 2188
QY 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
DB 2189 GGGAA-----TATTCTATTGGACAAGTACTCACTGTATTCTTTCTGTATTAAATTGG 2242
QY 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
DB 2243 GCTTTTAGTGTGGACAGGCATCTCCAAGCATTTGAAGCATTTGCCAAATGCAAGAGGACA 2302
QY 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
DB 2303 GCTTATGAATCTTCAAGATAATTGATATAGCCAAATTTGACAGCTATTCCGAAGAGT 2362
QY 341 GlyTyrLysProGluSerIleGlyThrValGluPheLysAsnValSerPheAsnTyr 360
DB 2363 GGGCACAAACCATTAATTAAGGAAATTTGGAATTTGCAAAATGTTCACTTCAGTTAC 2422
QY 361 ProSerArgProSerIleLysIleLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
DB 2423 CCATCTCGAAAAAGATTGAAGATCTTGAAGGGGCTTGAAGCTGAAGTGCAGAGTGGCAG 2482
QY 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
DB 2483 ACCTGGCTCTGGTGGAAACAGTGGCTGTGGAGAGACACACAGTCCAGCTGATGTCAG 2542
QY 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
DB 2543 AGGCTCTATGACCCACAGAGGGGATGGTCAGTGTGTGATGGACAGGATATTAGGACCAT 2602
QY 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
DB 2603 AATGTAAAGTCTTACGGGAAATCATTTGGTGTGTGAGTCAGGAACCTGTATTGTTGCC 2662
QY 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluMetGlu 460
DB 2663 ACCAGATAGCTGAAACCAATTCGTATGGCCGTGAATAATGTACCATGGATGATGATGAG 2722
QY 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
DB 2723 ARAAGCTCAAGGAAGCAATGCTTATGATCTTATCATGAACCTGCTCATAAATTTGAC 2782
QY 481 ThrLeuValGlyLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
DB 2783 ACCCTGGTGGAGAGAGGGGCCCAATTTGAGTGTGGGAGAGAGAGGATTCGCCATT 2842
QY 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
DB 2843 GCACGTGCCCTGGTTCGCAACCCCAAGATCTCTCTCTGGATGAGGCCACGTGAGCCTTG 2902
QY 521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
DB 2903 GACACAGAAAGCGAAGCAGTGGTTTCAGTGGCTCTCGATAAG 2944
RESULT 13

; Sequence 3, Application US/09769097
; Patent No. US20020055128A1
; GENERAL INFORMATION:
; APPLICANT: Kimberly Anne Brun
; APPLICANT: Richard James Chenery
; APPLICANT: Harma Ellens
; APPLICANT: John Anthony Feild
; APPLICANT: Lin Yue

; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES

; FILE OF INVENTION: ENCODING RAT MDRIA AND SCREENING METHODS THEREOF

; FILE REFERENCE: GP-50009-C2

; CURRENT APPLICATION NUMBER: US/09769, 097

; CURRENT FILING DATE: 2001-01-25

; PRIOR APPLICATION NUMBER: 09/208, 809

; PRIOR FILING DATE: 1998-12-09

; PRIOR APPLICATION NUMBER: 09/156, 800

; PRIOR FILING DATE: 1999-09-17

; PRIOR APPLICATION NUMBER: US99/20770

; PRIOR FILING DATE: 1999-09-10

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 3

; LENGTH: 4425

; TYPE: DNA

; ORGANISM: RATTUS RATTUS

US-09-769-097-3

Alignment Scores:

Pred. No.:	1.24e-157	Length:	4425
Score:	1543.00	Matches:	300
Percent Similarity:	73.04%	Conservative:	101
Best Local Similarity:	54.64%	Mismatches:	102
Query Match:	56.33%	Indels:	46
DB:	10	Gaps:	5

US-09-873-409-7 (1-541) x US-09-769-097-3 (1-4425)

Qy	1	MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu	20
Db	499	ATGTCGTGGGAACCTGGCGGCATTTATCCATGGAATGGCTCCCACTTATGATGCTG	558
Qy	21	ValLeuGlyLeuMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr	40
Db	559	GTCTTTGGAGACATGACAGATAGCTTTGCCAAAT-----GTAGGAACAACCGTAGTATG	612
Qy	41	SerPheArgLeuThr-----	46
Db	613	AGTTTCTACAATGTACAGACATATATGCCAAGCTGGAGGACGAAATGACCACTAGCC	672
Qy	47	LeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeu	66
Db	673	TACTATTACAGGCGCATTTGGCGGTGCTCATCTGCTTACATCCAGTTTCCACT	732
Qy	67	TriPleIleThrAlaAlaArgGlnThrLysArgIleArgLysGlnPheHisSerVal	86
Db	733	TGGTGCTGGCAGCTGGGAGACAAATACACAAGATTAGGCAGAGTTTTCATGCCATC	792
Qy	87	LeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyLeuLeuAsnThrArgMet	106
Db	793	ATGAATCAGGAGATAGGCTGGTTTGGAGTGTGATGAGTTTACTCGCGCTGGAAGCTAACT	852
Qy	107	Thr---AspIleAspIleSerAspGlyIleGlyAspIleAlaLeuLeuPheGln	125
Db	853	ACAGATGACGCTCCAAATTAATGAGGAATTTGGTACAAATTTGAAATGTTCTTTTCAG	912
Qy	126	AsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLysGlyTyrLysLeuThr	145
Db	913	GCAATGGCAACATTTTGGTGGTTTATATAGATTACTCGCGCTGGAAGCTAACT	972
Qy	146	LeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAlaCysSerArgMet	165
Db	973	CTTGTGATTTTGGCCATCAGCCCTCTTGGAGTGTGATTTGGGCAAGATA	1032

Qy	166	ValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGlu	185
Db	1033	TTGTCTTTTACTGATAAGGAACCTCAGGCTTATGCAAAAGCTGGAGCAGTTGCTGAA	1092
Qy	186	GluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGlnLysGluLeuGln	205
Db	1093	GAAGTCTTAGCAGCCATCAGAACTGTGATTGCTTGGAGGACAAAAGAGGAACTTGA	1152
Qy	206	ArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeu	225
Db	1152	-----	1152
Qy	226	SerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly	245
Db	1153	-----AGGTACAATAACAATTTGGAAGAGCTAAAGGCTTGGG	1191
Qy	246	IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGly	265
Db	1192	ATAAAGAAAGCTATACGGCCCAACATTTCCATGGGTGCAGCTTTTCTGCTTATCTATGCA	1251
Qy	266	ThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr	285
Db	1252	TCATATGCTCTGGCATTCTGGTATGGGACTTCTTGGTCAATCTCAAAAGAA-----TAC	1305
Qy	286	ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly	305
Db	1306	ACTATTGGACAAGTGTCTCACTGTCTTTTCTGTATTAAATGGAGCAATTCAGTGTGGG	1365
Qy	306	AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe	325
Db	1366	CAGGCATCTCCAATATTGAAGCTTTCGCCAATGCTAGAGGAGCAGCTTATGAAGTCTTC	1425
Qy	326	GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu	345
Db	1426	AGTATAATTGATAAATAAGGCCAGTAGACAGCTTCTCAAAAGAGTGGGCAACCCGAC	1485
Qy	346	SerIleGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer	365
Db	1486	AACATACAGGAATTTGGAAATTCAAAATATTCACCTTCAGTTACCCGCTCGAAAGAC	1545
Qy	366	IleLysIleLeuLysGlyLeuAsnLeuArgLysSerGlyGluThrValAlaLeuVal	385
Db	1546	GTTTCAGATCTTGAAGGCGCTCAACCTGAAGGTGAAGAGCGGCGAGCAGTAGCCTGGTT	1605
Qy	386	GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro	405
Db	1606	GGCAACAGTGGCTGTGGGAAAGACAACTGTCTCAGCTGCTGAGAGGCTCTACGACCCC	1665
Qy	406	AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr	425
Db	1666	ATAGAGGCGGAGTCACTATCGACGAGCAGGACATCAGGACCATCAATGTGAGGTATCTG	1725
Qy	426	ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn	445
Db	1726	CGGGAATCATTTGGGTGGTGTGAGTCAGGAACCGGTGTGTGTTCGCCACCAATTCGCCAA	1785
Qy	446	AsnIleLysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGlu	465
Db	1786	AACATTCGCTATCGCGAGAAACGTCACCATGGATGAGTAGAGAAAGCTGTCAAGAA	1845
Qy	466	AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu	485
Db	1846	GCCAAATGCTATGATTTTCATGAAACTGCCCCACAAATTTGACACCCCTGGTGGTGGAG	1905
Qy	486	LysGlyAlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal	505
Db	1906	AGAGGCGCGCAGCTGAGTGGGGGACAGAAACAGAGGATGCCCATTCGCCGGCCTGGTC	1965
Qy	506	ArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys	525
Db	1966	CGAACCCCAAGATCTCTTTTGGTGGAGGCGCACGCTGAGCCTTGGACACAGAAAGCGAA	2025
Qy	526	SerAlaValGlnAlaLeuGluLys	534

Db 2026 GCCGTGTTTCAGCGCGCTCTGGATAAG 2052
::: |||||||

RESULT 14

US-09-769-097-1

; Sequence 1, Application US/09769097
; Patent No., US20020055128A1
; GENERAL INFORMATION:
; APPLICANT: Kimberly Anne Brun
; APPLICANT: Richard James Chenery
; APPLICANT: Harna Ellens
; APPLICANT: John Anthony Feild
; APPLICANT: Lin Yue

; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES
; FILE OF INVENTION: ENCODING RAT MDRIA AND SCREENING METHODS THEREOF
; FILE REFERENCE: GP-50009-C3
; CURRENT APPLICATION NUMBER: US/09/769, 097
; CURRENT FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 09/208, 809
; PRIOR FILING DATE: 1998-12-09
; PRIOR APPLICATION NUMBER: 09/156, 800
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: US99/20770
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 4369
; TYPE: DNA
; ORGANISM: RATTUS RATTUS
US-09-769-097-1

Alignment Scores:
Pred. No.: Length: 2,016-157 4369
Score: 1541.00 Matches: 300
Percent Similarity: 73.04% Conservative: 101
Best Local Similarity: 54.64% Mismatches: 102
Query Match: 56.26% Indels: 46
DB: 10 Gaps: 5

US-09-873-409-7 (1-541) x US-09-769-097-1 (1-4369)

Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 499 ATGCTGCTGGAACTCTGGCGGCATATCCATGGAAATTCCTCCACTTATGATGCTG 558
Qy 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
Db 559 GTCTTTGGACATGACATAGCTTTTGCAAT- - - - -GTAGGAAACACCCGTAGTATG 612
Qy 41 SerPhePheArgLeuThr- - - - - 46
Db 613 AGTTTCTACAAATGCTACAGACATATATGCCAAGCTGGAGACGAAATGGCCAGTACGCC 672
Qy 47 LeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeu 66
Db 673 TACTATTACGGGCAATGTTGGCGGTGCTCATCGTCCCTACATCCAGGTTTCATT 732
Qy 67 TrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGlnPhePheHisSerVal 86
Db 733 TGTGCTGTCAGCTGGGACACAAATACACAGATTAGGCAGAAATTTTCCATGCCATC 792
Qy 87 LeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGluLeuAsnThrArgMet 106
Db 793 ATGAATCAGGAGATAGGCTGTTTGACGTGCATGACGTTCGGGAGCTCAACACCCGGCTC 852
Qy 107 Thr- - -AspIleAspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGln 125
Db 853 ACAGATGACGCTCCCAAAATTAATGAAGGAATTTGGTGACAAATTTGGAATGTTCTTTCAG 912
Qy 126 AsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLysGlyTyrPheLysLeuThr 145
Db 913 GCAATGCCAACATTTTTTGTGGTTTTTATAATAGGATTTACTCGCGGCTGGAAAGCTAACT 972

Qy 146 LeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAlaAalaCysSerArgMet 165
Db 973 CTTGTGATTTTGGCCATCAGCCCTCTCTTGACCTGTAGCTGGTATTTGGGCAAGATA 1032
Qy 166 ValIleSerLeuThrSerLysGluLeuSerLysAlaGlyAlaValAlaGlu 185
Db 1033 TTGCTTCTTACTGATAAGGAATCCAGGCTTATGCAAAAGCTGGAGCAGTTGCTGAA 1092
Qy 186 GluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGlnLysGluLeuGln 205
Db 1093 GAAGTCTTAGCAGCCATCAGAACTGTGATTCCTTTGGAGGACAAAAGAGCACTTGAA 1152
Qy 206 ArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeu 225
Db 1152 - - - - - 1152
Qy 226 SerCysValLeu- - - - -PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 245
Db 1153 - - - - -AGGTACAATAAATTTGGAAGAGCTAAAGGCTTGGG 1191
Qy 246 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGly 265
Db 1192 ATAAAGAAAGCTATCAGGCCCAACATTTCCATGGGTGGAGCTTTCTGCTTATCTATGCA 1251
Qy 266 ThrTyrGlyLeuAlaPheThrTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 285
Db 1252 TCATATGCTCTGGCATTCGTGATGGACTTCTCTTGGTCTCATCTCAAAAGAA- - - - -TAC 1305
Qy 286 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 305
Db 1306 ACTATTGGACAAGTCTCACTGTCTTTTCTGTATTATTTGGAGCATTCAGTGTGG 1365
Qy 306 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 325
Db 1366 CAGGCATCTCCAAATATTGAAGCTTCGCAATGTAGAGGAGCAGCTTATGAAGTCTTC 1425
Qy 326 GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 345
Db 1426 AGTATAATTGATAAATAAGCCAGTATAGACAGCTTCTCAAGAGAGTGGGCACAAACCCGAC 1485
Qy 346 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 365
Db 1486 AACATACAGAAATTTGGAATTTCAAAAATATTCACTTCAGTTACCCGCTCGAAAGAC 1545
Qy 366 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 385
Db 1546 GTTCAGATCTTGAGGCGCTCAACCTGAAGTGAAGAGCGGCGGACGAGCGTAGCCCTGTT 1605
Qy 386 GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro 405
Db 1606 GGCAACAGTGGCTGTGGGAAAAGCACAACTGTCAGCTGTCGAGAGGCTCTACGACCCC 1665
Qy 406 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 425
Db 1666 ATAGAGGGCGAGTCTAGTATCGACGACAGGACATCAGGACCATCAATGTGAGGTATCTG 1725
Qy 426 ArgAspHisIleGlyValSerGlnGluProValLeuPheGlyThrThrIleSerAsn 445
Db 1726 CGGAAATCATTTGGGTGGTGTAGTCAGGAACCCGCTGTTGCCACCAATTCGGGAA 1785
Qy 446 AsnIleLysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGlu 465
Db 1786 AACATTCGTTATGGCCGAGAAACGTCACCATGATGATAGAGAAGCTGTCAAGAA 1845
Qy 466 AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 485
Db 1846 GCCAATGCCTTATGATTTTCATGAAACTGCCCCCAAAATTTGACACCTCTGTTGTGTAG 1905
Qy 486 LysGlyAlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal 505
Db 1906 AGAGGGCGCAGCTAGTGGGGGACAGAAACAGAGATCGCCATTGCCCGGCGCTTGGTC 1965


```
Qy 506 ArgAsnProLysIleLeuAlaLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 525
Db 1966 CGCAACCCCAAGATCCCTTTGTTGATGGCCACAGCTCAGCTTGGACACAGAAAGCGAA 2025

Qy 526 SerAlaValGlnAlaLeuGluLys 534
Db 2026 GCCGTGGTTCAGGCGCTCTGGATAAG 2052

RESULT 15
US-10-072-621-2
; Sequence 2, Application US/10072621
; Patent No. US20020169137A1
; GENERAL INFORMATION:
; APPLICANT: Reiner, Peter B.
; APPLICANT: Connop, Bruce P.
; APPLICANT: Pollard, Michelle
; TITLE OF INVENTION: REGULATION OF AMYLOID PRECURSOR PROTEIN EXPRESSION
; FILE REFERENCE: 100103.402
; CURRENT APPLICATION NUMBER: US/10/072,621
; CURRENT FILING DATE: 2002-02-08
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 4643
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-072-621-2

Alignment Scores:
Pred. No.: 3,65e-157 Length: 4643
Score: 1539.00 Matches: 305
Percent Similarity: 73.10% Conservative: 100
Best Local Similarity: 55.05% Mismatches: 99
Query Match: 56.19% Indels: 50
DB: 9 Gaps: 7

US-09-873-409-7 (1-541) x US-10-072-621-2 (1-4643)

Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 575 ATGGTGGTGGGAACCTTGGCTGCCATCATCCATGGGGCTGGACCTCTCTCATGATGCTG 634

Qy 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
Db 635 GTGTTTGGAGAAATGACAGATATCTTCCAAATGCAGAAATTTAGAGATCTGATGCA 694

Qy 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
Db 695 AACATCACTAATAGAGTATGATCAATACATACAGGGTTCTTCATGAATCTGGAGGAAGAC 754

Qy 44 -----ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db 755 ATGACCGATGATGCTATATTATACAGTGGAAATGGTGTGGGGTGTGGTGTGCTCTTAC 814

Qy 62 IleGlnIleSerLeuTyrIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db 815 ATTACAGTTTCAATTTGGTGTGGCTGCACCTGGAACACAAATACACAAATTTAGAAACAG 874

Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101
Db 875 TTTTTCATGCTATAATCGCAGCAGGATAGCTGGTTTGTATGTGCACGATGTTGGGGAG 934

Qy 102 LeuAsnThrArgMetThr-----AspIleAspIleSerAspGlyIleGlyAspIle 120
Db 935 CTTRACACCCGACTTACAGATGATGCTCTCCAAAGATTAATGAAGGAATTTGGTGACAAAT 994

Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValClyLeuValLys 140
Db 995 GGAATGTTCTTTCAGTCAATGGCACATTTTTCATCGGTGTTATAGAGATTACAGT 1054

Qy 141 GlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db 2045 GCACGTGCCCTGGTTCGCAACCCCAAGATCTCTCTCTGGATGAGGCCACGCTCAGCCTTG 2104
```

```
Db 1055 GGTGGAAGCTAACCCCTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTGCTGCTGCT 1114
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db 1115 GTCTGGCAAGATACCTTCTTCTTACTGTATAAAGAACTCTTAGCGTATGCAGAAAGCT 1174
Qy 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db 1175 GGAGCAGTAGCTGAAGAGGCTTGGCAGCAATTAGAACTGTGATTGCATTGGAGGACAA 1234
Qy 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTyrPheTyrPhe 220
Db 1235 AAGAAAGAACTTCAA----- 1249
Qy 221 ProGlnTyrLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
Db 1250 -----AGGTACAACAATAATTTAGAGAA 1273
Qy 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
Db 1274 GCTAAAGAAATTTGGATAAAGAAAGCTATTACAGCCAATATTCTATAGTGTGCTCTTC 1333
Qy 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerIleLeuLeuAsn 280
Db 1334 CTGCTGATCTATGCATCTTATGCTCTGGCCTTCTGGTATGGACCACTTGGTCTCTCA 1393
Qy 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
Db 1394 GGGAA-----TATCTATTTGACAAAGTACTCAGTGA---TTTTCTGATTAATTGGG 1444
Qy 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
Db 1445 GCTTTTAGTGTGGACAGGCATCTCCAAAGCATTTGAAGCATTTGCAAAATGTCAGAGGAC 1504
Qy 321 AlaPheHisIlePheGlnValIleAspLysProSerIleAspAsnPheSerThrAla 340
Db 1505 GCTTATGAAATCTTCAAGATAATTTGATAAGCAAGTATGACACATTTCCGAAGAGT 1564
Qy 341 GlyTyrLysProGluSerIleGlyThrValGluPheLysAsnValSerPheAsnTyr 360
Db 1565 GGGCACAAACAGATATATTAAGGGAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 1624
Qy 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
Db 1625 CCATCTCGAAAGAAAGTAAAGATCTTCAAGGGTCTGAACCTGGAAGTGCAGAGTGGGCAG 1684
Qy 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
Db 1685 ACGTGGCCCTGGTTGGAAACAGTGGCTGTGGGAAGACACACAGTCCAGCTGATGCG 1744
Qy 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
Db 1745 AGGCTCTATGACCCACAGAGGGATGTCAGTGTGATGGACAGAGATATTAGGACCATA 1804
Qy 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
Db 1805 AATGTAAGTCTTACCGGAAATCATTTGGTGTGGTGTGAGTCAGGAACCTGATTGTTGCC 1864
Qy 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspValThrAspGluGluMetGlu 460
Db 1865 ACCACGATAGCTGAAACATTCGCTATGGCCGTGAAATGTCACCATGGATGAGATGAG 1924
Qy 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
Db 1925 AAAGCTCTCAAGAGCAATGCTATGACTTTATCATGAAGTCCCTCATAAATTTGAC 1984
Qy 481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
Db 1985 ACCCTGTTGGAGAGAGAGGGGCCCTGAGTTGAGTGGTGGCAGAGAGAGGATGCCATT 2044
Qy 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
Db 2045 GCACGTGCCCTGGTTCGCAACCCCAAGATCTCTCTCTGGATGAGGCCACGCTCAGCCTTG 2104
```


QY 521 AppSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
||:::|||||:::||||| ||||| |||||:::|
Db 2105 GACACAGAAAGCGAGCAGTGGTTCAGGTGGCTCTGGATAAG 2146

Search completed: March 31, 2003, 16:14:26
Job time : 261.92 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4 p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 30, 2003, 01:24:25 ; Search time 7313.73 Seconds
(without alignments)
2152.747 Million cell updates/sec

Title: US-09-873-409-7

Perfect score: 2739

Sequence: 1 MILGILASLVNGACPLMLP.....SEKSAVQALEKDPTRYSF 541

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -MODEL=frame+ p2n model -DEV=xlh

-Q=/cgn2_1/USPTO_spool/US09873409/runat_27032003.115418.19228/app_query.fasta_1.7544
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09873409 @CGN 1 1 54887 @runat_27032003.115418.19228 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl :
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.ste.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*

29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sv.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2737	99.9	2021	6	AX339033 Sequence
2	2697	98.5	3702	6	AX339031 Sequence
3	2568.5	93.8	1940	6	AX339034 Sequence
4	2528.5	92.3	3621	6	AX339032 Sequence
5	2437	89.0	3699	6	AX478104 Sequence
6	1868	68.2	3177	6	AX339030 Sequence
7	1566.5	57.2	4189	6	AX322791 Sequence
8	1566.5	57.2	4298	10	MUSMDR
9	1561.5	57.0	4279	6	AX105082 Sequence
10	1557.5	56.9	4279	6	AX105078 Sequence
11	1557.5	56.9	4279	6	AX105080 Sequence
12	1557	56.8	4195	6	AX108656 Sequence
13	1556.5	56.8	4045	12	AF269224
14	1555.5	56.8	3860	6	AX322787
15	1555.5	56.8	4378	6	E02326
16	1555.5	56.8	4646	6	AX391099
17	1555	56.8	4280	10	CRUPGPII
18	1552.5	56.7	3888	10	AY082609
19	1552.5	56.7	4233	6	AR123273
20	1552.5	56.7	4279	6	AX105057
21	1552.5	56.7	4317	4	AF045016
22	1552.5	56.7	4317	6	AX105059
23	1552.5	56.7	4669	6	AR091275
24	1552.5	56.7	4669	6	AR203322
25	1548.5	56.5	4186	6	AX108654
26	1548	56.5	4254	6	AX401748
27	1548	56.5	4254	10	RATMDRM
28	1548	56.5	4927	10	AF257746
29	1547	56.5	4323	10	AF286167
30	1546.5	56.5	3860	6	AX322789
31	1546.5	56.5	4646	6	AX336420
32	1546.5	56.5	4646	6	AX336708
33	1546.5	56.5	4646	6	I49610
34	1546.5	56.5	4646	9	HUMMDR1
35	1546.5	56.5	4669	6	I08557
36	1546.5	56.5	6505	6	AR028671
37	1546.5	56.5	8630	6	AX012320
38	1546.5	56.5	8630	6	AX012321
39	1546.5	56.5	9318	6	AR028672
40	1545	56.4	4264	6	AR051647
41	1545	56.4	4264	6	AR051650
42	1542	56.3	4356	10	MUSMDRAA
43	1541.5	56.3	4296	10	CRUPGPI
44	1540.5	56.2	4304	10	CRUPGPI165
45	1540	56.2	3489	4	AB029153

ALIGNMENTS

RESULT 1

AX339033
LOCUS AX339033 2021 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 15 from Patent WO0194400.
ACCESSION AX339033
VERSION AX339033.1 GI:18129125
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Frank, M.H. and Sayegh, M.H.
TITLE A gene encoding a multidrug resistance human p-glycoprotein
homologue on chromosome 7p15-21 and uses thereof
JOURNAL Patent: WO 0194400-A 15 13-DEC-2001;
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES
Location/Qualifiers
1..2021
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 547 a 406 c 462 g 605 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 2,61e-214 Length: 2021
Score: 2737.00 Matches: 541
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.93% Indels: 0
DB: 6 Gaps: 0
US-09-873-409-7 (1-541) x AX339033 (1-2021)
QY 1 MetLeuGlyLeuLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
DB 34 ATGATCTGGGTATACCTGGATCACTGGTCAATGGAGCGCTTCCTTTAAATGCCACTG 93
QY 21 ValLeuGlyGluMetSerAspAsnLeuLeuSerGlyCysLeuValGlnThrAsnThrTyr 40
DB 94 GTTTTAGAGAAATAGTGAATTAACCTTTATAGTGGATGCTTAGTCCAAACTAACACATAC 153
QY 41 SerPhePheArgLeuThrLeuTyrValGlyLeuValAlaLeuLeuPheGly 60
DB 154 TCCTTCTTCAGGTTGACCCCTGATATATGTTGGAATAGCTGTGCTGCTTGAATTTTGGT 213
QY 61 TyrLeuGlnLeuSerLeuTyrPheLeuAlaArgGlnThrLysArgLeuArgLys 80
DB 214 TACATACAGATTTCTCTGTGGATTATTAACCTGCACGACGACGACGAGGATTCGAAAA 273
QY 81 GlnPhePheHisSerValLeuAlaGlnAspLeuGlyTyrPheAspSerCysAspLeuGly 100
DB 274 CAGTTTTTTTCATTTCAGTTTGGCACAGACATCGCTGGTTGTATAGCTGTGACATCGGT 333
QY 101 GluLeuAsnThrArgMetThrAspLeuAspLysLeuSerAspGlyLeuGlyAspLysLeu 120
DB 334 GAACCTTAACACTCGCATGCACAGACATTCACAAATCAGTCATGATGATTCGAGATAAGATT 393
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerLeuGlyLeuAlaValGlyLeuValLys 140
DB 394 GCTCTGTGTTTCAAAAACATGCTACTTTTCGATTTGGCTGGCAGTTGTTGGTGAAG 453
QY 141 GlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuLeuMetAlaSerAlaAla 160
DB 454 GGCTGGAAACTCACCTTAGTGACTCTATCCACGCTCTCTCTTATATAGGCTTCAGCGGCA 513
QY 161 AlaCysSerArgMetValLeuSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
DB 514 GCATGTTCTAGGATGGTCACTCTCATTCACAGTAAGGAATTAAGTGCCTATTCCAAAGCT 573
QY 181 GlyAlaValAlaGluValLeuSerSerLeuArgThrValLeuAlaPheArgAlaGln 200
DB 574 GGGGCTGTGCAGAGAAGTCTTGTCTATCAATCCGAAACAGTCATAGCTTTTAGGGCCAC 633

QY 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnLeuThrArgTyrAlaTyrPheTyrPhe 220
DB 634 GAGAAAGAACTTCAAGGCTCTTCTTTTAAATAATAACAAGATATGCTTTGGTTTATTTT 693
QY 221 ProGlnTyrPheLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
DB 694 CCCAGTGGCTACTAGTTGTGTCTGTTTNTTGTAAAGGTATACACAGAATCTCAAGAT 753
QY 241 AlaLysAspPheGlyLeuLysArgThrLeuAlaSerLysValSerLeuGlyAlaValTyr 260
DB 754 GCAAAGGATTTGGCATAAAAGGACTATAGCTTCAAAAGTGTCTCTTGGTGTCTGTATC 813
QY 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuLeuAsn 280
DB 814 TTCTTTATGAATGAACCTATGACTTGTCTTTTGTATGAACCTCTCTTGAATCTTAAT 873
QY 281 GlyGluProGlyTyrThrLeuGlyThrValLeuAlaValPhePheSerValLeuHisSer 300
DB 874 GGAGAACTGGATATACCATCGGACTGTCTTGTCTGCTTCTTTAGTGTAAATCCATAGC 933
QY 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
DB 934 AGTTATGCTATTGGAGCAGCAGTCCCTCACTTTGAACCTTCGCAATAGCCGAGAGCT 993
QY 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
DB 994 GCCTTTTCATATTTCCAGCTTATTGATAAGAAACCCAGCATATAGATAACTTTCCACAGCT 1053
QY 341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
DB 1054 GGATATAAACCTGAATCCATAGAGGAACCTGTGGAAATTTAAATAATGTTCTTTCAATTAT 1113
QY 361 ProSerArgProSerIleLysLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
DB 1114 CCATCAAGACCATCTATCAAGATTTGAAAGGTCTGAAATCTCAGAAATTAAGCTCGAGAG 1173
QY 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
DB 1174 ACAGTCGCTTGGTGGTCTCAATGSCAGTGGGAGAGTACGCTAGTCTCCAGCTTCTGCAG 1233
QY 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
DB 1234 AGGTTATATGATCCGATGATGGCTTTATCATGGTGGATGAGAATGACATCAGAGCTTTA 1293
QY 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
DB 1294 AATGCGCGCATTAUCGACACCATATGAGTGGTGTAGTCAAGAGCTGTTTGTTCGGG 1353
QY 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspValThrAspGluGluMetGlu 460
DB 1354 ACCACCATCAGTAACAAATATCAAGTATGACGAGATGATGTGACTGATGAAGAGATGGAG 1413
QY 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
DB 1414 AGAGCAGCAAGGAGCAAAATGCTATGATTTTATCATGGAGTTTCTTAATTAATTAAT 1473
QY 481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
DB 1474 ACATTGGTAGGGGAAAAGAGCTCAAAATGAGTGGAGGCGAGAAACAGAGGATTCGCAAT 1533
QY 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
DB 1534 GCTCGTGCCTTAGTTTCGAAACCCCAAGATTTCTGATTTTAGATGAGGCTAGCTCTGCCCTG 1593
QY 521 AspSerGluSerLysSerAlaValGlnAlaLeuGluLysAspThrProArgTyrSer 540
DB 1594 GATTCAAGAAAGCAAGTCAAGTGTTCAGCTGCTCAAGCTGACATGAGAGGATACCCCAAGTATTC 1653
QY 541 Phe 541
DB 1654 TTT 1656
RESULT 2

AX339031 AX339031 3702 bp DNA linear PAT 09-JAN-2002
LOCUS Sequence 13 from Patent WO0194400.
DEFINITION AX339031
ACCESSION AX339031
VERSION AX339031.1 GI:18129123
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE Frank, M.H. and Sayegh, M.H.
AUTHORS
TITLE A gene encoding a multidrug resistance human p-glycoprotein
homologue on chromosome 7p15-21 and uses thereof
JOURNAL Patent: WO 0194400-A 13 13-DEC-2001;
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES
Location/Qualifiers
source 1..3702
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 1098 a 708 c 856 g 1039 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 1,04e-210 Length: 3702
Score: 2697.00 Matches: 534
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.47% Indels: 0
DB: 6 Gaps: 0
US-09-873-409-7 (1-541) x AX339031 (1-3702)
Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 34 ATGATCCTGGGTATACCTGGATCCTGGTCAATGGAGCCCTGCTTCTTAATGCCACATG 93
Qy 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
Db 94 GTTTTAGGAGAAATGAGTATACCTTATTAGTGGATGCTAGTCCAAACTAACACATAC 153
Qy 41 SerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGly 60
Db 154 TCTTTCTTCAGGTGACCTCTGATATATGTTGGAATAGGTGTTGCTGCTGATTTTGGT 213
Qy 61 TyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLys 80
Db 214 TACATACAGATTTCTTGTGGATTATAACTGACAGCAGCAGACCAAGAGGATTCGAAA 273
Qy 81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGly 100
Db 274 CAGTTTTTTCATTCAGTTTGGCACAGGACATCGGCTGGTTGATAGCTGTGACATCGGT 333
Qy 101 GluLeuAsnThrArgMetThrAspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db 334 GAACCTTAACACTCCGATGACAGACATTCAGAAAATCAGTGTATGAGATAAGATT 393
Qy 121 AlaLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db 394 GCTCTGTGTTTCANAAACATGCTACTTTTCATTTGGCTGGCAGTTGGTTGGTGAAG 453
Qy 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db 454 GGCTGGAAACTACCTTAGTACTCTATCCACGCTCCTCTTATATAGGCTTCAGCGGCA 513
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db 514 GCATGTTCTAGGATGGTCATCTCATTGACCAGTAAGGAATTAAGTGCCTATTCCAAAGCT 573
Qy 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db 574 GGGGCTGTGGCAGAGAAGTCTTGTATCATCAATCCGAACAGTCATAGCTTTTAGGCCCCAG

Qy 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
Db 634 GAGAAAGAACTTCAAAGGTCTTTCTTTAAATATAACAAGATATGCTTGGTGTATTTT 693
Qy 221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
Db 694 CCCAGTGGCTACTAAGTTGTGTTCTGTTNTTGTAGGTATACACAGAACTCAAGAT 753
Qy 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
Db 754 GCAAGAGATTTTGGCATAAAGGACTATAGCTTCAAAAGTGTCTCTGGTGTGTGTAC 813
Qy 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
Db 814 TTCTTTATGAATGGAAACCTATAGACTTGTCTTGTCTGTTCTTTAGTGTATTCCTTAAT 873
Qy 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
Db 874 GGAGAACCTGGATATACCATCGGACTGTTCTTGTCTGTTCTTTAGTGTATTCATAGC 933
Qy 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
Db 934 AGTTATTGCTATGGAGCAGCAGTCCCTCACTTTGAAACCTTCGCAATAGCCGAGGAGCT 993
Qy 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
Db 994 GCCTTTTCATATTTCCAGGTTATTGATAGAAACCCAGTATAGATAACTTTTCCACAGCT 1053
Qy 341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
Db 1054 GGATATAAACCTGAATCCATAGAGAACTGTGGAAATTTAAAAATGTTCTTCTTCAATTAT 1113
Qy 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
Db 1114 CCATCAAGACCATCTATCAAGATTTCTGAAAGGTCTGAATCTCAGAATTAAGTCGAGAG 1173
Qy 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
Db 1174 ACAGTCGCCCTTGGCTGCTCAATGCGCAGTGGGAAAGATACGGTAGTCCAGCTTCTGCAG 1233
Qy 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
Db 1234 AGTTATATGATCCGATGATGCTTTATCATGGTGGATGAGAAATGACATCAGAGCTTTA 1293
Qy 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
Db 1294 AATGTCGCGCATTAATCGACACCATATTTGAGTGGTTAGTCAAGAGCCTGTTTGTTCGGG 1353
Qy 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
Db 1354 ACCACCATCAGTAACAATATCAAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAG 1413
Qy 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
Db 1414 AGACAGCAAGGGAAGCAAAATGCGTATGATTTTATCATGGAGTTTCTTAATAAATTAAT 1473
Qy 481 ThrLeuValGlyLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
Db 1474 ACATGTTAGGGGAAAAAGAGCTCAATGAGTGGAGGCGCAGAAACAGAGGATCGCAATT 1533
Qy 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
Db 1534 GCTCGTGCCTTAGTTCCAAACCCCAAGATTTCTGATTTTAGATGAGGCTACGTCTGCCCTG 1593
Qy 521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
Db 1594 GATTCAAGAAAGCAAGTCAGCTGTTCAAGTCGACTGGAGAAG 1635
RESULT 3
AX339034 AX339034 1940 bp DNA linear PAT 09-JAN-2002
LOCUS Sequence 16 from Patent WO0194400.
DEFINITION AX339034
ACCESSION AX339034

VERSION AX339034.1 GI:18129126
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1
JOURNAL Frank, M.H. and Sayegh, M.H.
FEATURES A gene encoding a multidrug resistance human p-glycoprotein
source homologue on chromosome 7p15-21 and uses thereof
1. 1940 Patent: WO 0194400-A 16 13-DEC-2001;
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 530 a 394 c 448 g 568 t
ORIGIN

Alignment Scores:
Pred. No.: 1,45e-200 Length: 1940
Score: 2568.50 Matches: 514
Percent Similarity: 95.01% Conservative: 0
Best Local Similarity: 95.01% Mismatches: 0
Query Match: 93.78% Indels: 27
DB: 6 Gaps: 1

US-09-873-409-7 (1-541) x AX339034 (1-1940)

QY 1 MetLeuGlyLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
DB 34 ATGATCTGGGTATCTAGTGGCATCACTGGTCAATGGAGCTGCCTTCCTTTAAATGCCACTG 93
QY 21 ValLeuGlyGluMetSerAspAsnLeuLeuSerGlyCysLeuValGlnThrAsnThrTyr 40
DB 94 GTTTTAGGAGAAATGAGTGATTAACCTTATTAGTGGATGTCTAGTCCAACTAAACACATAC 153
QY 41 SerPhePheGluThrLeuTyrTyrValGlyIleGlyValAlaLeuLeuPheGly 60
DB 154 TCYTCTTCAGTGTACCTGTATTATGTGGATAGTGTGTCTGCTGTGATTTTGGT 213
QY 61 TyrIleGlnLeuSerLeuTyrIleThrAlaAlaArgGlnThrLysArgIleArgLys 80
DB 214 TACATACAGATTTCTTGTGGATTATAAATGCAGCACGACAGCAAGAGGATTCGAAAA 273
QY 81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGly 100
DB 274 CAGTTTTTTCATTCAGTTTTTGGCACAGGACATCGGCTGGTTTGATAGCTGTGACATCGGT 333
QY 101 GluLeuAsnThrArgMetThrAspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
DB 334 GAACCTTAACACTCGCATGACAGACATTTGACAAATCAGTGATGGTATTGGAGATAGATT 393
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
DB 394 GCTCTGTGTTTCAAAACATGTCTACTTTTTCGATTGGCTGGCAGTTGGTTGGTGAAG 453
QY 141 GlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuLeuMetAlaSerAlaAla 160
DB 454 GGCTGGAAACTCACCTAGTAGTACTATCCACAGTCTCCTCTTATAATGGCTTCAGCGGCA 513
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
DB 514 GCATGTTCTAGGATGGTCACTCATCTATTGACCCAGTAGGAAATTAAGTGCCTATTCCAAAGCT 573
QY 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
DB 574 GGGCTGTGGCAGAGAAGTCTTGTGCATCAATCCGAACAGTCATAGCTTTTAGGGCCAG 633
QY 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
DB 634 GAGAAAGAACTTCAA----- 648

QY 221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
DB 649 -----AGGTATACACAGAAATCTCAAGAT 672
QY 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
DB 673 GCAAGAGATTTTGGCATAAAAGGACTATAGCTTCAAAAGTGTCCTTGGTGCTGTGTAC 732
QY 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuLeuLeuAsn 280
DB 733 TTTCTTTATGAATGAAGAACCTTAGACTTGTCTTTTGGTATGAACCTCTTGTATCTTAAT 792
QY 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
DB 793 GGAGAACCTGGATATACCATCGGACTGTCTTGTCTTTCTTTAGTGTAAATCCATAGC 852
QY 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
DB 853 AGTTATTGCAATTGGAGCAGCAGTCCCTCACTTTGAACCTTCGCAATAGCCGAGAGCT 912
QY 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
DB 913 GCCTTTTCATATTTTCCAGGTTATTGATAGAAACCCAGATATAGATAACTTTTCCAGCT 972
QY 341 GlyTyrLysProGluSerIleGlyThrValGluPheLysAsnValSerPheAsnTyr 360
DB 973 GGATATAAACCTGAATCCATCCATAGAGGAACCTGTGGAATTTAAATAATGTTCTTTCAATTAT 1032
QY 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
DB 1033 CCATCAAGACCATCTATCAAGATTCTCAAGAGGTCTCAATCTCAGAATTAAGTCTGGAGAG 1092
QY 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
DB 1093 ACAGTCCCTTGTCTCGTCTCAATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTTCGAG 1152
QY 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
DB 1153 AGGTTATATGATCCGGATGATGGCTTTATCATGGTGGATGAGAATGACATCAGAGCTTTA 1212
QY 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
DB 1213 AATGTCCGCATTTATCGAGACCATATTTGGAGTGGTTAGTCAAGAGCTGTTTGTTCGGG 1272
QY 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspValThrAspGluGluMetGlu 460
DB 1273 ACCACCATCAGTAACAAATATCAAGTATGGACGAGATGTGTGACTGATGAAGAGATGGAG 1332
QY 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
DB 1333 AGAGCAGCAAGGGAAGCAAAATGCTATGATTTTATCATGGAGTTTCTTAATAATTTAAT 1392
QY 481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
DB 1393 ACATTGGTGGGAAAAAGGAGCTCAAAATGAGTGGAGGGCAGAAACAGAGGATTCGCAATT 1452
QY 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
DB 1453 GCTGTGGCTTAGTTCGAACCCCAAGATTCTGATTTTAGTAGGCTAGCTGTGCCCTG 1512
QY 521 AspSerGluSerLysSerAlaValGlnAlaLeuGluLysAspThrProArgTyrSer 540
DB 1513 GATTCAGAAAGCAAGTCAGCTGTTCAAGCTGCATCGAGAGGATACCCCCAGAGTATTCA 1572
QY 541 Phe 541
DB 1573 TTT 1575

RESULT 4
AX339032 AX339032 3621 bp DNA linear PAT 09-JAN-2002
LOCUS Sequence 14 from Patent WO0194400.
DEFINITION
ACCESSION AX339032


```

VERSION  AX339032.1  GI:18129124
KEYWORDS
SOURCE   human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE    Frank, M.H. and Sayegh, M.H.
JOURNAL  A gene encoding a multidrug resistance human p-glycoprotein
        homologous to chromosome 7p15-21 and uses thereof
        Patent: WO 0194400-A 14 13-DEC-2001;
        THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES
        Location/Qualifiers
        source
        1..3621
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
BASE COUNT  1081 a 696 c 842 g 1002 t
ORIGIN
Alignment Scores:
Pred. No.:      5.89e-197      Length:      3621
Score:          2528.50      Matches:      507
Percent Similarity: 94.94%      Conservative: 0
Best Local Similarity: 94.94%      Mismatches: 0
Query Match:      92.31%      Indels:      27
DB:                6          Gaps:          1
US-09-873-409-7 (1-541) x AX339032 (1-3621)
QY 1 MetIleuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
DB 34 ATGATCTGGGTATACCTGGCATCACTGGTCAATGGAGCCCTTCCTTTAAATGCCACTG 93
QY 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
DB 94 GTTTTAGGAGAAATGAGTATACCTTATAGTGGATGCTAGTCCAACTAACACATAC 153
QY 41 SerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaLeuIlePheGly 60
DB 154 TCTTTCTCAGGTGACCTGTATATATGTTGAATAGGTGCTGCTTGAATTTTGGT 213
QY 61 TyrIleGlnIleSerLeuTyrIleThrAlaAlaArgGlnThrIleArgIleArgLys 80
DB 214 TACATACAGATTTCTCTGTGGATATATAACTGCAGCAGCAGACAGAGAGATTCGAAA 273
QY 81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGly 100
DB 274 CAGTTTTTTCATTCAGTTTTGGCACAGGACATCGGCTGGTTGATAGCTGTGACATCGGT 333
QY 101 GluLeuAsnThrArgMetThrAspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
DB 334 GAACCTTAACACTCGCATGACAGACATTTGACAAATCAGTCATGATGTTGGAGATAAGATT 393
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
DB 394 GCTCTGTGTTTTCAAACATGCTACTTTTTTCGATTTGGCCCTGGCAGTTGTTGGTGAAG 453
QY 141 GlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
DB 454 GGCTGGAACCTCACCTAGTAGTCTATCATCGCTCTCTTATATAGCTTTCAGCGGCA 513
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
DB 514 GCATGTTCTAGGATGGTCATCTCATTGACCAGTAAGGAATTAAGTGCCTATTCCAAAGCT 573
QY 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
DB 574 GGGGCTGTGGCAGAAGAAGTCTTGTATCAATCCGACAGTCATAGCTTTTAGGGCCACG 633
QY 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTyrPheTyrPhe 220
DB 634 GAGAAAGAACTTCA----- 648

```

RESULT 5

AX478104

LOCUS

DEFINITION

Accession

Version

Keywords

Source

Organism

AX478104

Sequence 34 from Patent WO0240541.

Accession AX478104

Version AX478104.1 GI:22217064

Keywords human.

Source

Organism Homo sapiens

linear

DNA

3699 bp

PAT 12-AUG-2002

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE AUTHORS

1. Tang, Y.T., Yue, H., Nguyen, D.B., Hafalia, A.J., Elliott, V.S., Lu, Y., Walla, N.K., Yao, M.G., Baughn, M.R., Gandhi, A.R., Ding, L., Sanjanwala, M., Rankumar, J., Arvizu, C., Gietzen, K.J., Lal, P.G., Azimzai, Y., Khan, F.A., Thangavelu, K., Thornton, M., Lu, D.A., Tribouley, C.M., Warren, B.A., Ison, C.H., Das, D., Raumann, B.E., Policky, J.L. and Kearney, L.

Transporters and ion channels

Patent: WO 0240541-A 34 23-MAY-2002;

Incyte Genomics, Inc. (US)

FEATURES

Location/Qualifiers

source

1. 3699

/organism="Homo sapiens"

/db_xref="taxon:9606"

/notes="Incyte ID No: 7472030CB1"

BASE COUNT 1116 a 707 c 860 g 1016 t

ORIGIN

Alignment Scores:
Pred. No.: 1.81e-189 Length: 3699
Score: 2437.00 Matches: 495
Percent Similarity: 90.71% Conservative: 3
Best Local Similarity: 90.16% Mismatches: 9
Query Match: 88.97% Indels: 42
DB: 6 Gaps: 3

US-09-873-409-7 (1-541) x AX478104 (1-3699)

Qy 1 MetileLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 145 ATGATCTGGGTACTGACATCACTGTTCAATGAGGCTGCTTCTTAAATGCCACTG 204
Qy 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
Db 205 TGTATAGAGAAATGAGTGAATACCTTATTAGTGGATGCTAGTCCACATCAACACA 264
Qy 41 SerPhePheArg-----LeuThr 46
Db 265 AATTATCAGAACTGACTCTCAAGAGAGCTGAATGAAGATATGACTCTGTTGACC 324
Qy 47 LeuTyrTrpValGlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeu 66
Db 325 CTGTATTATGTTGGAATAGGTGTTGCTGCTTGATTTTGGTTACATACAGATTTCTCTG 384
Qy 67 TrpIleLeuThrAlaAlaArgGlnThrLysArgIleArgLysGlnPhePheHisSerVal 86
Db 385 TGGATTATACTGCAGCAGCAGACCAAGAGGATTCGAAAACAGTTTTTTTCATTTCAGTT 444
Qy 87 LeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGluLeuAsnThrArgMet 106
Db 445 TTGGCAGAGGACATCGGCTGTTGATAGCTGTGACATCGGTGAACCTTAACATCGCATG 504
Qy 107 Thr---AspIleAspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGln 125
Db 505 ACAGATGACATTGACAAATCAGTGTGTTGAGATAGATGAGATAGATGCTCTGTTTCAA 564
Qy 126 AsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLysGlyTrpLysLeuThr 145
Db 565 AACATGTCATCTTTTTCGATTTGGCTGTCAGTTGTTGGTGAAGGGCTGGAACCTCACC 624
Qy 146 LeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAlaCysSerArgMet 165
Db 625 CTAGTGACTCTATCCAGTCTCCTCTTATAATGGCTTCAGCGGAGCATGTTCTAGGATG 684
Qy 166 ValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyValAlaValAlaGlu 185
Db 685 GTCATCTATTGACCAGTAAGGAATTAAGTGCCTATTCCAAAGCTGGGCTGTGGCAGAA 744
Qy 186 GluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGln 205
Db 745 GAAGTCTTGTTCATCAATCCGAACAGTCAATAGCTTTAGGCGCCAGGAGAAAGAACTTCAA 804

Qy 206 ArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeu 225
Db 804 ----- 804
Qy 226 SerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 245
Db 805 -----AGGTATACACAGAACTCTCAAGATGCAAGAGGATTTTGGC 843
Qy 246 IleLysArgThrIleAlaSerLysValSerLeuGlyValAlaValTyrPhePheMetAsnGly 265
Db 844 ATAAAAAGGACTATAGCTTCAAAAGTGTCTCTGGTGTCTGTACTTCTTTATGAATGGA 903
Qy 266 ThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 285
Db 904 ACCTATGGACTTCTTTTGGTATGAACCTCTCTGATTTCTTAATGAGAACCTGGATAT 963
Qy 286 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 305
Db 964 ACCATCGGACTGTTCTTCTGCTGTTTCTTTAGTGTATCCATAGTATGATTGCAATTGGA 1023
Qy 306 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 325
Db 1024 GCAGCAGTCCCTCACCTTTGAAACCTTCGCAATAGCCGAGAGCTGCTTTCAATTTTC 1083
Qy 326 GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 345
Db 1084 CAGTTATTGATAAGAAACCCAGTATAGTAACTTTTCCACAGCTGGATATAAACCTGAA 1143
Qy 346 SerIleGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 365
Db 1144 TCCATGAAGAACTGTGGAAATTTAAATAATGTTTCTTTCAATTATCCATCAAGGCCATCT 1203
Qy 366 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 385
Db 1204 ATCAAGATTTCTGAAAGGTCTGAAATCTCGGAATTAAGTCTGGAGAGACAGTCGCTTGC 1263
Qy 386 GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro 405
Db 1264 GGTCTCAATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAGAGTTTATATGATCCG 1323
Qy 406 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 425
Db 1324 GATGATGGCTTTTATCATGTTGATGAGATGAGATGACATCAGAGCTTTAAATGTGCGCATAT 1383
Qy 426 ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn 445
Db 1384 CGAGACCATATTGGAGTGGTGTAGTCAAGAGCTGTTTGTTCGGGACCACTCATCAGTAAC 1443
Qy 446 AsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGlu 465
Db 1444 AATATCAAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAGAGCAGCAAGGAA 1503
Qy 466 AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 485
Db 1504 GCAATTCGTATGATTTTATCATGGAGTTCCTAATAAATTAATACATTTGGTAGGGGAA 1563
Qy 486 LysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal 505
Db 1564 AAAGAGCTCAATAGTGGAGGCGAACAACAGAGGATCGCAATTTGCTCGCTTGTAGTT 1623
Qy 506 ArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 525
Db 1624 CGAAACCCCAAGATTTCTGATTTTAGATGAGCTACGCTGCTGCTGGATTCAGAAAGCAAG 1683
Qy 526 SerAlaValGlnAlaAlaLeuGluLys 534
Db 1684 TCAGCTGTTCAAGCTGCACCTGGAGAG 1710
RESULT 6
AX339030 3177 bp DNA linear PAT 09-JAN-2002
LOCUS
DEFINITION Sequence 12 from Patent WO0194400.

ACCESSION AX339030
VERSION AX339030.1 GI:18129122
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Frank, M.H. and Sayegh, M.H.
TITLE A gene encoding a multidrug resistance human p-glycoprotein
JOURNAL homologue on chromosome 7p15-21 and uses thereof
Patent: WO 0194400-A 12 13-DEC-2001;
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES
source
1. .3177
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 970 a 601 c 735 g 870 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 4,58e-143 Length: 3177
Score: 1868.00 Matches: 370
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 68.20% Indels: 0
DB: 6 Gaps: 0
US-09-873-409-7 (1-541) x AX339030 (1-3177)
QY 165 MetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAla 184
Db 1 ATGGTCATCTCATTACCAAGTAAGGTAAGTGCCTATTCCAAAGCTGGGCTGGCA 60
QY 185 GluGluValLeuSerSerIleAArgThrValIleAlaPheAArgLacGlnGluLysGluLeu 204
Db 61 GAAGAGTCTTGTTCATCAATCCCAACAGTCATAGCCCTTTAGGGCCCGAGGAGAAGACTT 120
QY 205 GlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeu 224
Db 121 CAAAGTCTTCTTTAAATATAACAGATATGCTTGTGTTTATTTTCCCCAGTGGCTA 180
QY 225 LeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPhe 244
Db 181 CTAAGTGTGTCTGTTTAAAGTATACACAGAACTCTCAAGATGCAAAAGGATTTT 240
QY 245 GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsn 264
Db 241 GGCATAAAAGGAGTATAGCTTCAAAAGTGTCTTGGTGTCTGTACTTCTTTATGAAT 300
QY 265 GlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGly 284
Db 301 GGAACTATGGACTTGTCTTTTGGTATGGAACTCTCTGTGTTCTTATGAGAACTTGG 360
QY 285 TyrThrIleGlyThrValLeuAlaValPheSerValIleHisSerSertyrCysIle 304
Db 361 TATACCATCGGACTGTCTTCTGTGTTTCTTTAGTGTATCCATAGCAGTATTGCATT 420
QY 305 GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 324
Db 421 GGAGCAGCAGCTCCCTCACTTTGAAACCTTCGCAATAGCCCGAGGAGTGCCTTCATATT 480
QY 325 PheGlnValIleAspLysValProSerIleAspAsnPheSerThrAlaGlyTyrLysPro 344
Db 481 TTCAGGTATTGATAGAAACCCAGTATAGATAAATTTTCCACAGCTGGATATAACCT 540
QY 345 GluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgPro 364
Db 541 GAATCCATAGAAGGAACTGTGGAAATTAAATAATGTTCTTTCATATTATCCATCAAGACCA 600
QY 365 SerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu 384
Db 601 TCTATCAAGATTCTGAAAGTCTGAATCTCAGAAATTAAGTCTGGAGAGACAGTCGCTTG 660

QY 385 ValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAsp 404
Db 661 GTCCGTCTCAATGGCAGTGGGAAGAGTACGCTAGTCCAGCTTCTGCAGAGTTATATGAT 720
QY 405 ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis 424
Db 721 CCGGATGATGGCTTTATCATGTGTGATGAGATGATCATCAGAGCTTTAAATGTGGCGCAT 780
QY 425 TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSer 444
Db 781 TATCAGAGACCATATTGGAGTGTAGTCAAGAGCCTGTTTTGTGGGACCACCATCAGT 840
QY 445 AsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArg 464
Db 841 AACAATATCAAGTATCGACGAGATGATGTGACTGATGAAGAGATGGAGAGCAGCAAG 900
QY 465 GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly 484
Db 901 GAAGCAATCGGTATGATTTTATCATGGAGTTCCTTAATAAATTTTATACATTGGTAGG 960
QY 485 GluLysGlyValAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu 504
Db 961 GAAAAGGAGCTCAATGATGTGGAGGCGAGAAACAGAGATCGCAATTCTCGTGCCTTA 1020
QY 505 ValArgAsnProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspSerGluSer 524
Db 1021 GTTCGAAACCCCAAGATTCTGATTTTAGATGAGGCTACGCTGCTCCCTGGATTTCAGAAAGC 1080
QY 525 LysSerAlaValGlnAlaLeuGluLys 534
Db 1081 AAGTCAGCTGTTCAAGCTGCCTGGAGAAG 1110
RESULT 7
AX322791 4189 bp DNA linear PAT 07-JAN-2002
LOCUS AX322791
DEFINITION Sequence 5 from Patent WO0192877.
ACCESSION AX322791
VERSION AX322791.1 GI:18093768
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Sorrentino, B. and Schuetz, J.
TITLE Method of identifying and/or isolating stem cells
JOURNAL Patent: WO 0192877-A 5 06-DEC-2001;
ST. JUDE CHILDREN'S RESEARCH HOSPITAL (US)
FEATURES
source
1. .4189
Location/Qualifiers
/organism="Mus musculus"
/db_xref="taxon:10090"
BASE COUNT 1204 a 875 c 1028 g 1082 t
ORIGIN
Alignment Scores:
Pred. No.: 2,78e-118 Length: 4189
Score: 1566.50 Matches: 307
Percent Similarity: 72.63% Conservatives: 99
Best Local Similarity: 54.92% Mismatches: 94
Query Match: 57.19% Indels: 59
DB: 6 Gaps: 6
US-09-873-409-7 (1-541) x AX322791 (1-4189)
QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 148 ATGATTCTGGAACTCTCGCTGCTATTATCCATGGAACATTACTTCCCTCTTGTGCTG 207
QY 21 ValLeuGlyGluMetSerAspAsnLeuIle----- 30
Db 208 GTGTTTGGAAACATGACAGATAGTTTTTACAAAAGCAGAACAGCCAGTATTCTGCAAGCAT 267


```
QY 31 -----SerGlyCysLeuValGlnThrAsnThr----- 39
Db 268 ACTAATCAAGTGGACCAACAGTACTCTGATCATCAGCAACAGCATCTCGGAGGAAGAG 327
QY 40 -----TyrSerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAla 56
Db 328 ATGCCATATACGCCTAC-----TATTACACCGGGATTGGTGGTGTG 372
QY 57 LeuIlePheGlyTyrIleGlnIleSerLeuThrIleIleThrAlaAlaArgGlnThrLys 76
Db 373 CTCATAGTTCCTACATCCAGGTTTCACTTTGGTCCCTCGGAGCTGGAACAGACATACAC 432
QY 77 ArgIleArgLysGlnPhePheHisPheValLeuAlaGlnAspIleGlyTyrPheAspSer 96
Db 433 AAGATTAGGAGAGATTTTCCATGCTATATGAATCAGGAGTAGCTGGTTGATGTG 492
QY 97 CysAspIleGlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGly 115
Db 493 CATGATGTTGGGAGCTCAACACCCGGCTCACAGATGATGCTCCAAAATTAATGACGGA 552
QY 116 IleGlyAspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAla 135
Db 553 ATTGGTCACAAAATTTGGGATGTTTTTTCAGTCCATACCAACATTTTAGCGGTTTTATC 612
QY 136 ValGlyLeuValLysGlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIle 155
Db 613 ATAGATTATTAAGTGGTTGGAGCTTAACCTTGTCATTTTGGCTGTGACGCCCTTTATT 672
QY 156 MetAlaSerAlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSer 175
Db 673 GGATTGTCATCTGCTTTGGGCAAGGATTTGATCTTCACTTACTAATAAGGAATCCAG 732
QY 176 AlaTyrSerLysAlaGlyAlaValAlaGluValLeuSerSerIleArgThrValIle 195
Db 733 GCTTATGCAAAAGCTGGAGCAGTGTGCTGAAGAAGTCTTAGCAGCCATCAGAACTGTGATT 792
QY 196 AlaPheArgAlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyr 215
Db 793 GCCTTTGGAGACACAGAGAACTTGAA----- 822
QY 216 AlaTrpPheTyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThr 235
Db 823 -----AGGTACAAT 831
QY 236 GlnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSer 255
Db 832 AAAAAATTTAGAAGAAGCTTAAATAATTTGGCATAAAGAAAGCTATCACAGCCAGCATTTGC 891
QY 256 LeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrTrpTyrGlyThr 275
Db 892 ATAGCATTCGCTACCTGTTGGTCTATGTCATCATATGCATCGGCATCTCTGGTATGGACA 951
QY 276 SerLeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePhe 295
Db 952 TCCTTGGTCTCTCAATGAA-----TATTCTATGGAGAGTCTTACTGCTCTCTTC 1005
QY 296 SerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAla 315
Db 1006 TCTATTTTGTTGGGACTTTTGTATTGAGACTTATGACACTTATGACATCGGCATCTCTGGTATGGACA 1065
QY 316 IleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAsp 335
Db 1066 AACGCACGAGGGGAGCCCTTTGAATCTTCAAGATAATTAATACGAGCCAGCATTTGAC 1125
QY 336 AsnPheSerThrAlaGlyTyrLysProGluSerIleGlyThrValGluPheLysAsn 355
Db 1126 AGCTTCTCAACAAAGGGCTACAAACACAGACAGTATATATGGGAAACTTAGAGCTTTAAAAAT 1185
QY 356 ValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArg 375
Db 1186 GTTCACTTCAACTACCCATCGAGAAGCAAGTTCAGATCTTGAAGGGCCCTCAATCTGAAG 1245
```

```
QY 376 IleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrVal 395
Db 1246 GTCAAGAGCGGACGACGCGTGGCTTTGGTTGGCAACAGTGGCTGTGTGAAAAAGCACAACT 1305
QY 396 ValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsn 415
Db 1306 GTCCAGCTGATGACAGAGGCTCTACGACCCCTCGAGGGCGTGTGATTCGACGACGACAA 1365
QY 416 AspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGlu 435
Db 1366 GACATCAGAACCATCATGTGAGGTATCTCAGGAGAGATCATTTGTTGGTGTGAGTACAGAA 1425
QY 436 ProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThr 455
Db 1426 CCTGTGCTGTTGCCACCATCGATCGCGGAGAACATTCGCTATGCGCGGAGAGATGTCAAC 1485
QY 456 AspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPhe 475
Db 1486 ATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAATGCCCTATCATCATGAACATG 1545
QY 476 ProAsnLysPheAsnThrLeuValGlyGluLysGlyValAlaGlnMetSerGlyGlyGlnLys 495
Db 1546 CCCACCAATTTGACACCTGTTGGTGTGAGAGAGGGCGGCGAGTGTGGGGACAGAAA 1605
QY 496 GlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGlu 515
Db 1606 CAGAGAATGCCATTGCGCGGCGCTGTCGCAATCCCAAGATCCTTTTGTGGACGAG 1665
QY 516 AlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
Db 1666 GCCACCTCAGCCCTGGATACAGAAAGTGAAGCTGTGTGTGCGAGCGGCACCTGGATAAG 1722

RESULT 8
MUSMDR      MUSMDR      4298 bp      mRNA      linear      ROD 04-OCT-1995
LOCUS      Mouse multidrug resistant protein mRNA, complete cds.
DEFINITION M14757.1 GI:199100
ACCESSION  M14757
VERSION    M14757.1
KEYWORDS   membrane-associated glycoprotein; multidrug resistance protein.
SOURCE     Mus musculus (strain BALB/c, sub_species domesticus) (clone: lambda
          DR11) cDNA to mRNA.
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 4298)
AUTHORS     Gros, P., Croop, J. and Housman, D.
TITLE       Mammalian multidrug resistance gene: complete cDNA sequence
            indicates strong homology to bacterial transport proteins
JOURNAL     Cell 47 (3), 371-380 (1996)
MEDLINE     87028229
PUBMED      3768958
COMMENT     The mdr gene family includes two or more related but distinct
            cellular genes.
            There are 3 possible start codons besides the one given in the
            Features table, at positions 161-163, 221-223 and 257-259. If the
            latter two start codons are used as initiator the first hydrophobic
            domain could possibly serve as a signal sequence.
FEATURES
  source
    1. .4298
      /organism="Mus musculus"
      /strain="BALB/c"
      /sub_species="domesticus"
      /db_xref="taxon:10090"
      /clone="lambda DR11"
      /cell_type="pre-B cell"
      <1. .4298
      /product="mdr mRNA"
      110. .3940
      /codon_start=1
      /product="multidrug resistance protein"
      /protein_id="AAA79005.1"
      /db_xref="GI:387426"
      /translation="MEFEENLKGKADKNFVKMGKKKKEKKKPAVGVFGMFRYADW"
  mRNA
  CDS
```


LDKLCMLTGTLAAI IHTGTLPLILMLVFGNWTDSPTTKAASLIPSI TWQSGPNSTLILS
NSFEDEEMALYIITDYGIGAVLIIVAIQVUSLWCAAGRQIHKIRKJFHFHAIINQGEJ
WVVDHDEYALNTDYGKINDGTGDKTGMFQSPITPTFLAGTIFIGISGWKTLVIL
AVSPILIGLSSALWAKVLTSTIKELQAKAGAAVEYLVAIRTVIAYFGQOQKELERY
NKULEEAKNVIKATVTSIGTILVYVYASLAWFYWSTSVLSNSESIGEVLTVP
PSILIGFTSIGHLANIEAFANARGAAPEI PKI INDEPS IDSPSTCKVPDS TGMNLE
PKVNFYPSRSEVQILKGLNLKVSGQTVLAVGNSGCKSTVTQMLQRLYDPLEGVV
SIDQDQITINVRILKIEGVSPQVTFATTAIIVIRIGREDVTQMBEAKAKEANA
YDFITMKLPQDFTLVGERAQIUSGQKQRIATARALVRNPKILDELSALDITSEEA
VQAALDQKREHRTTIVIAHRLTGVGNQDITAPDQGVIVBQGNHDELMREKGTYFKL
TQNTQBNETIPGNNAYGNSQSDTASELTSBESKSLPRLRSI SVRHKRQOQERRLS
KEATVEDPVLVSFWRLNLINSEMPYLVGLVINGVINGCQIOPVAFVPSRIVGVSRD
DDHETKQCNLISLFLVWGLISIVFYTFQFGTQKAGEILTGRVRYMFKMLRQD
ISWFDHNKSTGSLTTLSDASSVKGAMGARLAVTQNVANGLQVILSLVYQWLT
LLVYVILPLIGLITIMKLQJOKDKRQLEISGKTAEIENPRTIVSLVTRQOK
ETMYAQVLPYRNAMKKAHFVITPSFTQMMYFVSECFRFGALVLAQMLTFNV
MLVPSVVGMAAGNESTSFAPADYAKASASHIIRIETKPEIDISTYSTGLKPTLLE
GNVFPNGVQVNYPTRNI PVLOGLSLSEVKKGTOTALVSSGCGKSTVOLLERPYDPM
AGSVFLHGTIKQLNQMWRUHLIGVYBQPLPFDCSIENATAYGDNRSVRHEIYVRA
AKSEYHOFIDISLPQKYNTRVDKQGLQSGGQKQRIATARALVQPHILLLDEBSATSL
DTSEKVOYBALDQKBEHRTCTIVIAHRLSTIQNADLIVVINGVKVKGHGHQQLLAQK
GIYFSEVQAQAKES"

BASE COUNT	1226 a	903 c	1059 g	1110 t
ORIGIN				

Alignment Scores:

Pred. No.:	2,886-118	Length:	4298
Score:	1566.50	Matches:	307
Percent Similarity:	72.63%	Conservative:	99
Best Local Similarity:	54.92%	Mismatches:	94
Query Match:	57.19%	Indels:	59
DB:	10	Gaps:	6

US-09-873-409-7 (1-541) x MUSMDR (1-4298)

Qy	1	MetIleuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu	20
Db	257	ATGATTCTGGGA	316
Qy	21	ValLeuGlyGluMetSerAspAsnLeuLeu	30
Db	317	GTGTTTGGAAACATGACAGATAGT	376
Qy	31	-----SerGlyCysLeuValGlnThrAsnThr	39
Db	377	ACTAATCAAAGTGACCAACAGTACTCTGATCATCAGCAACAGCAGTCTGGAGGAAGAG	436
Qy	40	-----TyrSerPhePheArgLeuThrLeuTyrValGlyIleGlyValalaLa	56
Db	437	ATGGCCATATACGCTAC	481
Qy	57	LeuIlePheGlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLys	76
Db	482	CTCATAGTTGCTACATCCAGGTTTCACTTTGGTGCCTGGCAGCTGGAAGACAGATACAC	541
Qy	77	ArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSer	96
Db	542	AAGATTAGCCAGAAGTTTTTCCATGCTTATTAATCAGGAGATAGGCTGGTTTGATGTG	601
Qy	97	CysAspIleGlyGluLeuAsnThrArgMetThr	115
Db	602	CATGATGTTGGGGAGCTCAACACCCGGCTCACATGATGTCTCCAAAAATTAATGACGGA	661
Qy	116	IleGlyAspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAla	135
Db	662	ATTGGTGACAAAATGGGATGTTTTTTCAGTCCATAACACATTTTTTAGCCGGTTTTATC	721
Qy	136	ValGlyLeuValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIle	155
Db	722	ATAGGATTTATAAGTGGTTGGAAGCTAACCCCTTGTCATTTTTGGCTGTGCAGCCCTCTATT	781
Qy	156	MetAlaSerAlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSer	175

Db	782	GGATTGTCATCGCTTTGGTGGCCAAAGGTATTGCACTTCATTACTTAATAAGGAACCTCCAG	841
Qy	176	AlaTyrSerLysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIle	195
Db	842	GCTTATGCCAAAGCTGGAGCAGTTGCTGAAGAAGTCTTAGCAGCCATCAGAACTGTGATT	901
Qy	196	AlaPheArgAlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyr	215
Db	902	GCCTTTGGAGGACAAACAGAAAGGAACCTTGAA	931
Qy	216	AlaIrpPheTyrPheProGlnIrpLeuLeuSerCysValLeu***PheValArgTyrThr	235
Db	932	-----AGGTACAAT	940
Qy	236	GlnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSer	255
Db	941	AAAAATTTAGAAAGCTAAAAATGTTGCCATAAAGAAGCTATCACAGCCAGCATTCG	1000
Qy	256	LeuClAlaValTyrPheMetAsnGlyThrTyrGlyLeuAlaPheTyrGlyThr	275
Db	1001	ATAGGCATGCCCTACTGTTGGTCTATGCATCATATGCACCTGGCATCTCGTATGGGACA	1060
Qy	276	SerLeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePhe	295
Db	1061	TCCTTGGTCTCTCAATGNA-----TATTCTATTGGGAAGTCTTACTGCTCTCTTC	1114
Qy	296	SerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAla	315
Db	1115	TCATTTTGTGGGACATTTTAGTATTGCACTTTGGCCCAAAACATAGAAGCCTTTGCA	1174
Qy	316	IleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAsp	335
Db	1175	AACCCAGCGGGGCGCCTTTGAAATCTTCAAGATAATTGATAACGACCAAGCATTTGAC	1234
Qy	336	AsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsn	355
Db	1235	AGCTTCTCAACAAAGGCTACAACACAGACAGTATAATCGGGAACCTTAGAGTTTAAAAAT	1294
Qy	356	ValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArg	375
Db	1295	GTTCACTTCAACTACCATCGAAGCGGAAGTTTCAGATCTTGAGGGCCTCAATCTGAAG	1354
Qy	376	IleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrVal	395
Db	1355	GTGAAGACGGCAGACGGTGGCTTGGTTCGCAACAGTGGCTGTGGAAAAAGCACAACT	1414
Qy	396	ValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsn	415
Db	1415	GTCCAGCTGATGCAGAGGCTCTACGACCCCTGGAGGGCGGTGCTCAGTATCGACGGACAA	1474
Qy	416	AspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGlu	435
Db	1475	GACATCAGAACCATCAATGTGAGGTATCTGAGGGAGATCATTTGGTGTGGTCAAGCA	1534
Qy	436	ProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspValThr	455
Db	1535	CCTGTGCTGTTGCCACCGATTCGCCGAGAACATTCGCTATGGCCGAGAGAGATGTCAC	1594
Qy	456	AspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPhe	475
Db	1595	ATGATCGAGATTGAGAAAGCTCTCAGGAAGGCCAATGCCATGACTTCATCATGAAACTG	1654
Qy	476	ProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLys	495
Db	1655	CCCCACCAATTTGACACCCCTGGTTGTTGAGAGAGGGGGCGCAGTCTAGTGGGGGACAGAA	1714
Qy	496	GlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGlu	515
Db	1715	CAGAGAATCGCCATTTGCCGGGCGCTTGGTCCGCAATCCCAAGATCTCTTTTGTGGACAG	1774
Qy	516	AlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys	534
Db	1775	GCCACCTCAGCCCTTGGATACAGAAAGTGAAGTGTGGTGGACGCCGACTGATGAAG	1831

RESULT 9
AX105082
LOCUS AX105082 4279 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 26 from Patent WO0123540.
ACCESSION AX105082
VERSION AX105082.1 GI:13921232
KEYWORDS
SOURCE dog.
ORGANISM Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE 1 (bases 1 to 4279)
AUTHORS Stocker, P.J., Steimel-Crepi, D.T., Crepi, C.L., Reif, T.C. and
Patten, C.J.
TITLE P-glycoproteins and uses thereof
JOURNAL Patent: WO 0123540-A 26 05-APR-2001;
GENTEST CORPORATION (US)
FEATURES
Location/Qualifiers
1..4279
/organism="Canis familiaris"
/db_xref="taxon:9615"
17..3862
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC37723.1"
/db_xref="GI:13921233"
/translation="MDPEGRKGSAEKNFWMGKKKKKKKKKPTVSTFAMFRYSN
WDLRLNMLVGTMAAITHGAALPLMLVFNMTDSFANAGISRNKTPFVINESITNNT
QHPINLHEEMTYAYVYSGIGAGVLAAYIOVSFWLAAGQILKIRKOFFHAINRQ
EIGFWDVHDVGLNLTDDVSKINGEDIGKIMFFQSIATFTGIVGTFGRWKLTL
VIAAISVLGLSAIAWAKLSFTDKELLAYAKAGAVEVLAAIRTVIAFGOKKEL
ERYKNLEAKGIGIKATANIISIGNAFLIYASVALAWYCTSVLSSEVTIGOV
TVFVLIGAFSGQSPSIEAFANRGAAYEIPKIDNKPISDSISKSGHKPDNKG
NLFFKNVHFSYRKEKILKLNKLVQSGQTVLVNCGSGKSTTVQMLQRLYDPTD
GMVCDIQDIRTNVRLREITGVSQEPVLFATTAENIRYGRNVTWDEIBAVKE
ANAYDFIMKLPNKFDTLVGRGAQLSGQKQRIARALVNPKIALLDEATBALDTE
SEAVOVADKAKGRRTTIVIAHRLSTVRNADVIAGDDGVIKEGNHDELMEKGIY
PKLVOTGRNEVPELENATGESKESDALEMSPKDSGSLIKERSTERSHAPQGD
KLGTKEDLNENPVEFVFNWFLVGLIISFTFPGFTFGKAGEILTKRLRYNVPFRSM
LRQVSWDFDKNNFTALTRLANDAAQVKAIGSLAVITONANLGTGIIISLYG
WOLTLALLAIVPIAIAVGVEMKMLSGQKQRIARALVNPKIALLDEATBALDTE
EOKFEYMAQSLQVYNSLRKAHIFGVSPSITQAMMYFSYACGFRGAYLVANEFN
FQDVLVFSIAIVGMAVQVSSFPADYAKAKVAAHVMIEKSLPIDSYSPLGLKP
NTLEGNVFNVEVFNPTREDIPVIGLSLEVKKGQTLALVGSCKGCTVQQLERF
YDPLAGSVLIDGKEIKHLNWLRAHLGIYSQEPILFDCSIAENIAYGNSRVVSHDE
IVQAACEANHTHFIETLPEKYNTRVGDGKQQLSGQKQRIARALVNPKIALLDEA
TSALDTESEKVVQVQAGAKR"
LAQKGIYFSVMVQAGAKR"

BASE COUNT 1296 a 833 c 1009 g 1141 t
ORIGIN
Alignment Scores:
Pred. No.: 7,32e-118 Length: 4279
Score: 1561.50 Matches: 304
Percent Similarity: 72.12% Conservative: 97
Best Local Similarity: 54.68% Mismatches: 104
Query Match: 57.01% Indels: 51
DB: 6 Gaps: 5

US-09-873-409-7 (1-541) x AX105082 (1-4279)
Qy 1 MetLeuGlyLeuLeuAserLeuValAanGlyAlaCysLeuProLeuMetProLeu 20
Db 167 ATGTTGGGGGACAAATGGCTCATCATCCATGAGCTGCACCTCCCTCTCATGATGCTG 226
Qy 21 ValLeuGlyGluMetSerAspAsnLeuLeuSerGlyCysLeu----- 34
Db 227 GTTTTGGGAAACATGACAGATAGCTTTCGAATGACAGAAATTCAGAAACAAACCTTTT 286
Qy 35 -----ValGlnThrAsnThrTyrSerPhePheArg----- 44

Db 287 CCAGTTATAAATAATGAAAGTATTACGAAACAATACACAAATTCATCAACCATCTGGAG 346
Qy 45 -----LeuThrLeuTyrValGlyLeuValAlaLeuLeuPhe 59
Db 347 GAGGAATGACACAGTATGCTATTATTACAGTGGATCGTCTGGCTGGCTGGCTGGCT 406
Qy 60 GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg 79
Db 407 GCTTACATCCAGGTTTCATCTGGCTGGCAGCAGGAGACAGATACATCAAAATTAGA 466
Qy 80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99
Db 467 AAACAATTTTTCATGCTATCATCGACGAGAGATGGCTGGTTCAGCTGCAATGAGCT 526
Qy 100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyVsp 118
Db 527 GGGGAGCTTAACACCGGCTCACAGAGATGCTCCAAATCAATGAAGAAATTTGGCCAC 586
Qy 119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
Db 587 AAATTTGGAATGTTCTTCAATCAATAGCAACATTTTACCGGTTTATAGTGGGGTTT 646
Qy 139 ValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
Db 647 ACACGTGGTGGAGCTTAACCTTGTGATTTGGCCATCAGCCCTGTTCTTGGACATTTCA 706
Qy 159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
Db 707 GCGCCCATCTGGGCAAGATACATCTCTTATCTGATAAAGAACTCTTGGCTATGCA 766
Qy 179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198
Db 767 AAAGCTGGACAGTAGCTGAAGAAGCTCTTAGCAGCAATCAGAACTGTGATGCTTTGGA 826
Qy 199 AlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPhe 218
Db 827 GGACAAAGAAAGAACTTGAA----- 847
Qy 219 TyrPheProGlnTrpLeuLeuSerCysValLeu**PheValArgTyrThrGlnAsnLeu 238
Db 848 -----AGGTACAAACAAAAATTTA 865
Qy 239 LysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAla 258
Db 866 GAAGAAGCTTAAGGAATTTGGATAAGAAAGCTATCACGGCCCAACATTTCTATTGGTGCC 925
Qy 259 ValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIle 278
Db 926 GCTTTCTTATTGATCTATGTCATCATATGCTCTGGCTTTCTGGTATGGGACCTCTCTGTC 985
Qy 279 LeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIle 298
Db 986 CTCTCCAGTGAA-----TATACTATTGGACAGTACTCATCTGTCCTCTTTCTGTATTA 1039
Qy 299 HisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArg 318
Db 1040 ATTTGGGCTTTAGTATTGGACAGCATCCCAAGCATTTGAAGCATTTGCAACGCAAGA 1099
Qy 319 GlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSer 338
Db 1100 GGAGCAGCTATGAAATCTTCAAGATAATTTGACAATAATTAACCAAGCATTGACAGCTATTCG 1159
Qy 339 ThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPhe 358
Db 1160 AAGAGTGGACATAAACAGATAATTAATTAAGGAAATTTGAAATTTCAAAATTTGCTTCT 1219
Qy 359 AsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSer 378
Db 1220 AGTTACCTCTTCGAAAGAAAGTTAAGATCTTAAAGGCTCTCAACCTGAAGGTTTTCAGAGT 1279
Qy 379 GlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeu 398
Db 1280 GGGCAGACAGTGGGCTGGTGGGAAAGCAGTGGCTGCGGGAAGACGACGCGTGGAGCTG 1339

Qy 259 ValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIle 278
Db 926 GCTTCTTATTGATCATATGCTCTGCTTCTGATGAGCACTCTCTGCTC 985
Qy 279 LeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValile 298
Db 986 CTCCTCCAGTGAA-----TATTCATTGGACAAGTACTCACTGCTCTCTCTCTGTTATTA 1039
Qy 299 HisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArg 318
Db 1040 ATTGGGCTTTTAGTATTGGACAGCATCCCAAGCATTTGAAGCATTTGCAACAGCGAAGA 1099
Qy 319 GlyAlaAlaPheHisIlePheGluValIleAspLysLysProSerIleAspAsnPheSer 338
Db 1100 GGACAGCTTATGAAATCTTCAAGATTAATGACAATAACCAAGCATTTGACAGCTATTCG 1159
Qy 339 ThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPhe 358
Db 1160 AAGAGTGGACATAAACACAGATAATATTAAAGGAAATTTGGAATTCAAAAATGTTCACTTC 1219
Qy 359 AsnTyrProSerArgProSerIleIleLysIleLysGlyLeuAsnLeuArgIleLysSer 378
Db 1220 AGTTACCCCTCTCGAAAAGAGTTTAAGATCTTAAAGGCTCTCAACCTCAAGGTTACAGAT 1279
Qy 379 GlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGluLeu 398
Db 1280 GGGCAGACAGTGGCGCTGTTGGGAAACAGTGGCTCGGGAGACACACCGTGCACCTG 1339
Qy 399 LeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArg 418
Db 1340 ATGCAGAGGCTCTATGACCCACAGATGGCATGCTGTTGATGATGACAGGACATTTAGG 1399
Qy 419 AlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeu 438
Db 1400 ACCATAAATGTAAGGCATCTTCGGGAAATTAAGTCTGTTGTTGATGATGAGGAGCTGTTG 1459
Qy 439 PheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspValThrAspGluGlu 458
Db 1460 TTTGCCACCAAGTACGTTGAAAACATTCGCTATGCGCCGAAAATGTCACCATGATGAG 1519
Qy 459 MetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLys 478
Db 1520 ATTGAGAAAGCTGTTAAGGAAGCAATGCTATGATTTATCATGAAGTACCTTAATAAA 1579
Qy 479 PheAsnThrLeuValGlyGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIle 498
Db 1580 TTTGACACCTCTGTTGGAGAGAGAGGGCCCGACGTGAGTGGTGACACAGAAACAGAGAATC 1639
Qy 499 AlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSer 518
Db 1640 GCCATTGCTCGGGCTGTTGTCGCAACCCCAAGATTTCTTCTGCTGGATGAGGCAACGTCA 1699
Qy 519 AlaLeuAspSerGluSerLysAlaValGlnAlaLeuGluLys 534
Db 1700 GCTCTGGACACTGAAGTGAAGCAGTGTTCAGGTGGCCCTGGATAAG 1747
RESULT 11
AX105080
LOCUS AX105080 4279 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 24 from Patent WO0123540.
ACCESSION AX105080
VERSION AX105080.1 GI:13921230
KEYWORDS
SOURCE dog.
ORGANISM Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE 1 (bases 1 to 4279)
AUTHORS Stocker,P.J., Steinel-Crespi,D.T., Crespi,C.L., Reif,T.C. and Patten,C.J.
TITLE P-glycoproteins and uses thereof
JOURNAL Patent: WO 0123540-A 24 05-APR-2001;

GENTEST CORPORATION (US)
Location/Qualifiers
1. 4279
/organism="Canis familiaris"
/db_xref="taxon:9615"
17. 3862
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC37722.1"
/db_xref="GI:13921231"
/translation="NDPEGKRGSKAEKFWKMGKSKKKEKKKPTVSTFAMFRYSN
WDRILMLVGMTAAIIHGAALPMMLVFGNMTDSFANAGISRNKTKLIRKQFTHAINRQ
EIGWDFVDHVDLNRLLDDVSGINEGDKIMFFQSIATFTTGTIVGTGKWLTL
VILATSPVLGLSAAIWKILTSFTDKELLAYAKAGAVAEVLAARTVLSAFGQKKEL
EYVKNLBEAKGIGIKKATANISGAAPLLIYASVALAFWGTSLVLSSEYSGOVL
TVFFSVLIGAFSIGQASPIEAFANAGNAYEIKIIDNKPSIDSVSKSGHPDNLKG
NLEFNVHFSYPSRNEKILGNLKVQSGQTVLVNCGSCSTTVQMQRLIDPDK
GMVCDGQDITINVRHLREITGVVSQEPVLFATTIAENIRYGRNVNWDIEKAVKE
ANAYDFIMKLPNKFDTLVGERGAQLSGGOKRIARALVRNPKILLDEATSAKTE
SEAVQVALDKARKGRTTIVIAHRLSTVRNADVIAGDDGVIVEKGNHDELMEKGIY
FKLVMTQVRNEIELENATGESKSDALEMSPKDSGLIKRSTRRSIIHAPOGODR
KLGTEDLNENVPVPSFWRLKLNSTEWPFYVGIIFCALINGSLQAPFSIIISRIIGI
FRDEDPETKRONNMFSVLFLVGLIISFITFFLQGTFTGKAGEILUTKKLRITWFRSM
LRQVSWDFDKNTTCALTFLANDAAQVKAIGSLRVAITQNIANLGTGIIISLIY
WOLTLALLAIQPIIAIAGVEMKMLSGQALKDKKELEGAKIATEIENFRVVSILTR
EOKFEYMAQISQVPRNLRKAHIFGVFSITQAMVFSYAGCFRFGAYLVANEPMN
FQDVLVFSALVFGAMVQVSSFPADYAKAKVSAAHVIMIIKESPLIDSYSPHGLKP
NLEGNVTNEVFNPTRPDIPVLOGLSLEVKQGQTLALVSSSGGSKSTVQLLERF
YDPLAGSVLIDGKEIKHNLWRLHGVISQEPILFDCSIAENIAGYGNDSVSHHEE
IMQAAKEANIHFIEITPEKYNTRVGDGKTQSGGOKRIARALVRNPKILLDEA
TSALDTESEKVVQVEALDKAREGRTCTIVIAHRLSTIQNADLI VVFQNGKVKHGTHTQOL
LAQKGIYFSMVSVQAGAKH"
BASE COUNT 1296 a 833 c 1008 g 1142 t
ORIGIN
Alignment Scores:
Pred. No.: 1,55e-117 Length: 4279
Score: 1557.50 Matches: 303
Percent Similarity: 72.12% Conservative: 98
Best Local Similarity: 54.50% Mismatches: 104
Query Match: 56.86% Indels: 51
DB: 6 Gaps: 5
US-09-873-409-7 (1-541) x AX105080 (1-4279)
Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 167 ATGTTGGTGGGACATGCTGCCATCATCCATGGAGCTGCCTCTCTCATGATGCTG 226
Qy 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu----- 34
Db 227 GTTTTGGAAACATGACAGATAGCTTTGCAATGCAGGAATTTCAAGAAACAAACTTTT 286
Qy 35 -----ValGlnThrAsnThrTyrSerPhePheArg----- 44
Db 287 CCAGTTATAATTAAGAAAGTATTACGAACAATACACAACATTTCAACCACTCTGGAG 346
Qy 45 -----LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe 59
Db 347 GAGGAATGACCACGATGCTATTATTACAGTGGGATCGGTGCTGCGCTGCTGCTGCT 406
Qy 60 GlyTyrIleGlnIleSerLeuTyrIleThrAlaAlaArgGlnThrLysArgIleArg 79
Db 407 GCTTACATCAGGTTTTCATTCTGCTGCTGGCAGCAGAAAGACAGATCTCAAAATTAGA 466
Qy 80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIle 99
Db 467 AACAATTTTTCATGCTATCATCGCAGACAGAGATTGGCTGGTTTACGTGCTGACGCTT 526
Qy 100 GlyGluLeuAsnThrArgMetThr-----AspIleAspLysIleSerAspGlyIleGlyAsp 118
Db 527 GGGGAGCTTAACCCCGCTCACAGACGATGTCTCCAAATCAATGAAGGAATTTGGCGAC 586


```
QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 250 ATGGTGGTGGAACTTTGGCTGCATCATCCATGGAGCTGGAGTTCCTCTCATGATGCTG 309
QY 21 ValLeuGlyGluMetSerAspAsnLeuIleSer-----GlyCys 33
Db 310 GTGTTTGGAGCATGACGGATACCTTTGCAAAATGCAGAAATTTAGGAGATTTAGGAGCT 369
QY 34 LeuValGlnThrAenThrTyrSer----- 41
Db 370 CTGTGTTTAAACACACTAATAGCAGTAATATCACTGATCAGTGCCTGCATGAATCTG 429
QY 42 -----PhePheArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIle 58
Db 430 GAGGAAGATATGACCATGCTGCTATTAATACAGTGGAAATGGTGGCTGGCTGGTT 489
QY 59 PheGlyTyrIleGlnIleSerLeuThrIleIleThrAlaAlaArgGlnThrLysArgIle 78
Db 490 GCTGCTTACATTCAGGTTTCATTTTGGTGGCTGGCAGCTGGAAGACAAATACACAAAT 549
QY 79 ArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAsp 98
Db 550 AGAAAAAGTTCATGCTATTAATGCGCAGGAGATAGCTGGTTGATGTGCAGCAT 609
QY 99 IleGlyGluLeuAenThrArgMetThr---AspIleAspLysIleSerAspGlyIleGly 117
Db 610 GTTGGGAGCTTAACACCGGCTTACAGATGATGCTCCAAGATTAATGAAGGAATGGT 669
QY 118 AspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGly 137
Db 670 GACAAAATTTGGAATGTTCTTTTCACTCAATGCGCAACATTTTCACTGGGTTTATAGTAGGA 729
QY 138 LeuValLysGlyTyrPheLeuThrLeuValThrLeuSerThrSerProLeuIleMetAla 157
Db 730 TTACAGCTGGTGGAGCTGAACCTTGTGATTTGGCCATCAGTCTCTGTTCTTGGACTG 789
QY 158 SerAlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyr 177
Db 790 TCAGCTCAGCTGCGCAAGATACTGCTTCATTTACTGATAAAGAACTCTTAGCTTAT 849
QY 178 SerLysAlaGlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPhe 197
Db 850 GCAAAAGCTGGAGCAGTAGCTGAAGAGTCTTGGCAGCAATTAGAACTGTGATTCATTT 909
QY 198 ArgAlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTyr 217
Db 910 GGAGGACAAAAGAAAGAACTCGAA----- 933
QY 218 PheTyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsn 237
Db 934 -----AGGTACACAAAAT 948
QY 238 LeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGly 257
Db 949 TTAGAAGAACTAAAGAAATTTGGATAAGAAAGCTATTACAGCAATATTTCTTAGGT 1008
QY 258 AlaValTyrPheMetAsnGlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeu 277
Db 1009 GCTGCTTCTGCTTATCTATGATCATCTTATGCTCTGGCTCTGGTATGGAGCACCTTG 1068
QY 278 IleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerVal 297
Db 1069 GTCTCTCAAGAGNA-----TATTCATTGGACAGTACTCACTGATTTCTTTCTGTA 1122
QY 298 IleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAla 317
Db 1123 TTAATTTGGGCTTTTGTAGTTGGACAGGCTATCCCAAGCATTTGAAGCATTTGCAATGCA 1182
QY 318 ArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPhe 337
Db 1183 AGAGGAGCAGCTTTTGAATCTTCAAGATAATTTGATAATTAAGCAAGTATTGACAGCTAT 1242
```

```
QY 338 SerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSer 357
Db 1243 TCGAAGAGTGGGCAAAACCCAGATAATATTAAAGGAAATTTGGAATTCAGAAATTTTCAC 1302
QY 358 PheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLys 377
Db 1303 TTCAGTTTACCATTCTCGAAAAGAGTTAAAGATCTTGAAGGGCCTGAAACCTGAAAGGTGCA 1362
QY 378 SerGlyGlyThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGln 397
Db 1363 AGTGGGAGACGGTGGCCCTGGTTGGAACAGCGGCTGTGGAGAGACCAACGCTCCAG 1422
QY 398 LeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIle 417
Db 1423 CTGATGACAGAGGCTTTATGACCCACAGAGGGCATGCTCAGTCTGTGATGACAGGATATT 1482
QY 418 ArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProVal 437
Db 1483 AGGACCATAAACGTAAGGTTTCTTACGGGAATCATCGGTGTGTGAGTCAGGAACCTGTA 1542
QY 438 LeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspValThrAspGlu 457
Db 1543 TTGTTTCCACACAGATAGCTGAAACATTCGCTATGGTCTGTAAGATGTCACCATGGAT 1602
QY 458 GluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsn 477
Db 1603 GAGATTGAGAAAGCTGTCAAGGAAGCAATGCCTATGACTTTATCATGAAACTGCCTCAG 1662
QY 478 LysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArg 497
Db 1663 AAATTTGACACCTGTTGGAGAGAGAGGGGCCAGCTGAGTGTGTGGGCAAGACGAGG 1722
QY 498 IleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuAspGluAlaThr 517
Db 1723 ATGCCATTGACAGTGCCTCGTTTCGCAACCCCAAGATCTCTCTGTGGACGAGGCCAG 1782
QY 518 SerAlaLeuAspSerGluSerLysSerAlaValGlnAlaLeuGluLys 534
Db 1783 TCAGCCTTGGACACAGAAAGTGAAGCAGTGGTTTCAGGTGCTCTGGATAAG 1833
RESULT 13
AF269224 4045 bp DNA linear SYN 04-JUL-2000
LOCUS Synthetic construct Canis familiaris his-tagged-multidrug
DEFINITION resistance glycoprotein gene, complete cds.
ACCESSION AF269224
VERSION AF269224.1 GI:8926216
KEYWORDS synthetic construct.
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1 (bases 1 to 4045)
AUTHORS Roulet,A.C., Puel,O., Gesta,S., Alvinerie,M. and Pineau,T.J.
TITLE Analysis of dog MDR1 p-glycoprotein
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 4045)
AUTHORS Roulet,A.C., Puel,O., Gesta,S., Alvinerie,M. and Pineau,T.J.
DIRECT SUBMISSION
TITLE Submitted (19-MAY-2000) Pharmacology, INRA, 180 Chemin de
JOURNAL Tournefeuille, BP3, Toulouse 31931, France
FEATURES
source
1. .4045
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="original mRNA isolated from Beagle dog colon"
101. .3964
/codon_start=1
/transl_table=11
/product="his-tagged-multidrug resistance glycoprotein
MDR1"
/protein_id="AAF81747.1"
/db_xref="GI:8926217"
/translation="MHFHHDPEGGRKGSAEKFNFKMGKSKQNEKKEKPTVTSTFA
```


Qy	1	MetIleuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu	20
Db	269	ATGTTGGTGGGCAATGGGCTGCCATCATCGAGAGCTGCACCTCCCTCATGATGCTG	328
Qy	21	ValLeuGlyGluMetSerAspAsnLeuLeuSerGlyCysLeu	34
Db	329	GTTTTGGAAACATGACAGATAGCTTTGCCAAATGACAGGAATTTCAAGAAACAAACATTTT	388
Qy	35	ValGlnThrAsnThrTyrSerPheArg	44
Db	389	CCAGTTAATAATGAAGATGATTTACGAACATACACACATTTTCATCAACCATCTGGAG	448
Qy	45	LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe	59
Db	449	GAGGAATGACCCATGTCCTATTATTACATGGGATCGTGCTGGCGTGTGGTGGCT	508
Qy	60	GlyTyrIleGlnIleSerLeuTyrIleThrAlaAlaArgGlnThrIlysArgIleArg	79
Db	509	GCTTACATCCAGGTTTCATTCTGTGCTGTCGACGAGGAGACAGATACTCAAAATTAGA	568
Qy	80	LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle	99
Db	569	AAACAATTTTTTCATGCTATCATCGCAGAGAGATTGGCTGTTGCATGCGATGACGTT	628
Qy	100	GlyGluLeuAsnThrArgMetThr--AspIleAspIlyIleSerAspGlyIleGlyAsp	118
Db	629	GGGAGAGCTTAACCCCGGCTCACAGACGATGTCCTCCAAATCAATGAAGAAATGGCGAC	688
Qy	119	LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu	138
Db	689	AAATGTGAATGTTCTTTCAATCAATGACACATTTTTCACCGTTTTATAGTGGGGTTT	748
Qy	139	ValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer	158
Db	749	ACAGTGGTGGAGAGCTAACCCCTTGATTTGGCCATCAGCCCTGTTCTTGGACITTC	808
Qy	159	AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer	178

RESULT 14
AX322787
LOCUS AX322787 3860 bp DNA linear PAT 07-JAN-2002
DEFINITION Sequence 1 from Patent WO0192877.
ACCESSION AX322787
VERSION AX322787.1 GI:18093766
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Sorrentino B. and Schuetz, J.
TITLE Method of identifying and/or isolating stem cells
JOURNAL Patent: WO 0192877-A 1 08-DEC-2001;
ST. JUDE CHILDREN'S RESEARCH HOSPITAL (US)
FEATURES
source
1. 3860
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 1135 a 746 c 957 g 1022 t
ORIGIN
Alignment Scores:
Pred. No.: 1.99e-117 Length: 3860
Score: 1555.50 Matches: 306
Percent Similarity: 73.29% Conservative: 100
Best Local Similarity: 55.23% Mismatches: 99
Query Match: 56.79% Indels: 49
DB: 6 Gaps: 6
US-09-873-409-7 (1-541) x AX322787 (1-3860)
Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 151 ATGGTGGGAACTTTGGCTGCCATCATCCATGGGCTGGACTTCCTCTCATGATCGT 210
Qy 21 ValLeuGlyLeuMetSerAsp-----AsnLeuIleSer 31
Db 211 GTGTTTGGAGAAATGACAGATATCTTGGCAATCGAGAAATTTAGAAATCTGATGCA 270
Qy 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhe----- 43
Db 271 AACATCAATAAGATGATATCAATGATACAGGTTCTTCATGAATCTGGAGGAAGAC 330
Qy 44 -----ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db 331 ATGACCATATGCTTATTTACAGTGAATTTGGTGGGCTGGTGGTGGCTGCTTAC 390
Qy 62 IleGlnIleSerLeuTyrPheIleThrAlaAlaArgGlnThrLysArgIleAtgLYsGln 81
Db 391 ATTCAGGTTTCATTTGGCTGGAGCTGGAGAGCAATATACACAAATTTAGAAACAG 450
Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101
Db 451 TTTTTCATGCTATAATGCGACGAGATAGCTGGTTGATGTCACGATTTGGGGAG 510
Qy 102 LeuAsnThrArgMetThr-----AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db 511 CTTAAACCCGACTTACAGATGATGCTCTTAAGATTAATGAAGTATTGGTGACAAAATT 570
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db 571 GGAATGTTCTTTTCAGTCAATGGCAACATTTTTCATCTGGGTTTATAGTAGGATTTACACGT 630
Qy 141 GlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db 631 GTTGGAGACTAACCCCTGTGATTTGGCCATCAGTCTCTGTTCTTGGAGTCTCAGCTGCT 690
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db 691 GTCTGGGCAAGAGATACTATCTTCTTACTTATGATAAGAACTCTTAGCGTATGCAAAAGCT 750

Qy 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db 751 GGAGCAGTAGCTGAAGAGCTCTGGCAGCAATTAGAACTGTGATTGCATTTGGAGCACAA 810
Qy 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
Db 811 AAGAAAGAACTTGAA----- 825
Qy 221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
Db 826 -----AGGTACACAAATAATTAGAAAGAA 849
Qy 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
Db 850 GCTAAAAGAAATTTGGGATAAAGAAAGCTATTACAGCAATATTCTATAGTGCTGCTTTC 909
Qy 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsn 280
Db 910 CTGCTGATCTATGATCTTATGCTCTGGCTCTTCTGATGGGACCACTTGGTCTCTCA 969
Qy 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
Db 970 GGGGAA-----TATTCCTATTGGCAAGTACTCCTGATTTCTTTCTGATTAAATGGG 1023
Qy 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaAtgGlyAla 320
Db 1024 GCTTTTAGTGTGGACAGGATCTCCAAAGCATTTGAAGCATTTGCAAAATGTCAGAGGACA 1083
Qy 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
Db 1084 GCTTATCAAAATCTTCAAGATAATTGATAAGCAAGTATTGACAGCTATTTCGAAGAGT 1143
Qy 341 GlyTyrLysProGluSerIleGlyThrValGluPheLysAsnValSerPheAsnTyr 360
Db 1144 GGCACAAACCATGATATTAAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 1203
Qy 361 ProSerArgProSerIleLysIleLeuLysGlyLysAsnLeuArgIleLysSerGlyGlu 380
Db 1204 CCATCTCGAAAGAAAGTAAAGATCTTGAAGGGCTCTGAACCTGAAGGTGCAGAGTGGCAG 1263
Qy 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
Db 1264 ACAGTGGCTGCTGGTGGAAACATGCTGGTGGGAGAGACACACAGTCCAGCTGATGCAG 1323
Qy 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
Db 1324 AGCTCTATGACCCACAGAGGGGATGTCAGTGTGATGGACAGGATATTAGACCATA 1383
Qy 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
Db 1384 AATGTAAGGTTTCTACGGGAAATCATTTGGTGTGTCAGTCAGAACCTGATTGTTTGC 1443
Qy 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
Db 1444 ACCACGATGCTGAAACATTCGCTATGGCCGTGAAAAATGTCACCATGGATGAGATTGAG 1503
Qy 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
Db 1504 AAAGCTGTCAGGAAGCAATGCTATGACTTTATCATGAATGCTCTCAATAATTGAC 1563
Qy 481 ThrLeuValGlyGlyLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
Db 1564 ACCTGGTGGAGAGAGAGGGGCCAGTGTGAGTGTGGGAGAGAGAGATGCGCAT 1623
Qy 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
Db 1624 GCACGTGCCCTGGTGTGCAACCCCAAGATCTCTGCTGATGAGGCCACGTCAGCCCTG 1683
Qy 521 AspSerGluSerLysSerAlaValGlnAlaLeuGluLys 534
Db 1684 GACACAAAGCGAAGCAGTGGTTTCAGTGGCTCTCTGGATTAAG 1725

RESULT 15

E02326
LOCUS E02326 4378 bp RNA linear PAT 29-SEP-1997
DEFINITION Multidrug resistance relating gene derived from human normal cells.
ACCESSION E02326
VERSION E02326.1 GI:2170561
KEYWORDS JP 1990100680-A/1.
SOURCE Homo sapiens.
ORGANISM Homo sapiens

REFERENCE

1 (bases 1 to 4378)
Ueda,K. and Komano,T.
AUTHORS HUMAN NORMAL CELL-DERIVED MDR RELATED GENE
TITLE Patent: JP 1990100680-A 1 12-APR-1990;
JOURNAL SUNTORY LTD

COMMENT

OS Homo sapiens
PN JP 1990100680-A/1
PD 12-APR-1990
PF 05-OCT-1998 JP 1989251475
PI UEDA KAZUMITSU, KOMANO TORU
PC C12N15/12, C12N1/21, C12Q1/68;
CC strandedness: Single;
CC topology: Linear;
CC *source: tissue type=Adrenal gland;
CC *source: clone=SAM132;
FH Key Location/Qualifiers
FH 5'UTR 1..137
FT CDS 138..3980
FT /gene='Multidrug resistance relating gene' FT
FT mat_peptide 138..3977
FT /gene='Multidrug resistance relating gene' FT
FT 3'UTR 3981..>4378.

FEATURES

source
1..4378
/organism='Homo sapiens'
/db_xref='taxon:9606'
BASE COUNT 1315 a 818 c 1057 g 1188 t
ORIGIN

Alignment Scores:

Pred. No.: 2,33e-117 Length: 4378
Score: 1555.50 Matches: 306
Percent Similarity: 73.29% Conservative: 100
Best Local Similarity: 55.23% Mismatches: 99
Query Match: 56.79% Indels: 49
DB: 6 Gaps: 6

US-09-873-409-7 (1-541) x E02326 (1-4378)

Qy 1 MetileuGlyleuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 288 ATGGTGGTGGAACTTTGGCTGGCATCATCCATGGGGCTGGACTTCTCTCATGATGCTG 347
Qy 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
Db 348 GTGTTGGAGAAATGACAGATATCTTTGAAATGCGAGAAATTTAGAAATCTGATGCTCA 407
Qy 32 GlyCysLeuValGlnThr-----AsnThrTy+SerPhePhe----- 43
Db 408 AACATCACTAATAGAGTATATCAATGATACAGGGTTCTTCATGAATCTGGAGAGAC 467
Qy 44 -----ArgLeuThrLeuTyTyValGlyleGlyValAlaAlaLeuIlePheGlyTy 61
Db 468 ATGACAGGATGCTATTATTACAGTGGAAATGGTCTGGGGTGGTGGTCTGCTTAC 527
Qy 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db 528 ATTCAGGTTTCATTTTGGTGGCTGGAGCTGGAGCAAAATACACAAATTTAGAAACAG 587
Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101

Db 588 TTTTTCATGCTATATATGCGACAGGAGATAGGCTGGTTTGTGTGCACCATGTTGGGAG 647
Qy 102 LeuAsnThrArgMetThr--AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db 648 CTTACACCCCGACTTACAGATGCTCTCAAGATTAATGAAGGAATTTGGTACAAAT 707
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db 708 GGAATGTTCTTTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGGATTTACACGT 767
Qy 141 GlyTyLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db 768 GGTGGAGACTAACCTTGTGATTTTGGCCATCAGTCTGTCTTCTGGAGTGTGAGTGTCT 827
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db 828 GTCTGGGCAAGATACATCTTCAATTTACTGATAAGAACTCTTAGCGTATGCCAAGCT 887
Qy 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db 888 GGAGCAGTAGCTAGCTGAAGAGTCTTGGCAGCAATTAGAACTGTGATTGCAITTTGGAGACAA 947
Qy 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyRAlaTrpPheTyR 220
Db 948 AAGAAAGAACTTGAA----- 962
Qy 221 ProGlnTrpLeuLeuSerCysValLeu**PheValArgTyRThrGlnAsnLeuLysAsp 240
Db 963 -----AGGTACACAAATAATTAGAGAA 986
Qy 241 AlaLysAspPheGlyleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyR 260
Db 987 GCTAAAGAAATTTGGGATAAAGAAAGCTATTACAGCAATATTTCTATAGTGTGCTTTC 1046
Qy 261 PhePheMetAsnGlyTyRThrGlyLeuAlaPheTyRTrpTyRThrSerLeuIleLeuAsn 280
Db 1047 CTGCTGATCATGCTATGCTCTGGCTTCTGTGATGGGACCACTTGGTCTCTCA 1106
Qy 281 GlyGluProGlyTyRThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
Db 1107 GGGGAA-----TATCTATTGGCAAGTACTCCTGTTTCTGTTTCTGTTTAAATGGG 1160
Qy 301 SerTyrcysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
Db 1161 GCTTTTAGTGTGGACAGCATCTCCAGCATTTGAAGCATTTGCAAAATGCAAGAGAGCA 1220
Qy 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
Db 1221 GCTTATGAAATCTTCAAGATAATTGATAATAAGCCAAATTTGACACAGTATTGCAAGAGT 1280
Qy 341 GlyTyLysProGluSerIleGlyThrValGluPheLysAsnValSerPheAsnTyR 360
Db 1281 GGGCAAAACCAAGATATATTAAGGAAATTTGGAATTTGAGAAATGTTCACTTCAGTTAC 1340
Qy 361 ProSerArgProSerIleLysIleLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
Db 1341 CCATCTCGAAAGAAAGTTAAGATCTTGAAGGCTCTGAACTGAAAGTGCAGAGTGGGAG 1400
Qy 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
Db 1401 ACGGTGGCTGGTTGGAAACAGTGGCTGTGGGAAGAGACACACAGTCCAGCTGATGTCAG 1460
Qy 401 ArgLeuTyRAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
Db 1461 AGGCTCTATGACCCACAGAGGGGATGGTGGTGTGAGTGGAGCAGGATATTAGGACCATTA 1520
Qy 421 AsnValArgHisTyRArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
Db 1521 AATGTAAGGTTTCTACGGGAATCATTTGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1580
Qy 441 ThrThrIleSerAsnAsnIleLysTyRArgAspAspValThrAspGluGluMetGlu 460


```
Db 1581 ACCACGATAGCTGAAAACATTTCGCTATGGCCGTGAAATGTCACCATGGATGAGATTGAG 1640
QY 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
Db 1641 AAGCTGTCAGGAGAGCCATGCCATATGACTTTATCATGAACACTGCCTCATAAATTGAC 1700
QY 481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
Db 1701 ACCCTGGTTGGAGAGAGAGGGGCCCACTTGAGTGGTGGGCAGAGCAGAGGATCGCCATT 1760
QY 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
Db 1761 GCACGTGCCCTGGTTCGTAACCCCAAGATCCTCCTGCTGATGAGGCCACGTCAGCCTTG 1820
QY 521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
Db 1821 GACACAGAAAGCGAAGCAGTGGTTTCAGGTGGCTCTGGATAG 1862
```

Search completed: March 31, 2003, 02:23:23
Job time : 7369.73 secs

GenCore version 5.1.4 p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 30, 2003, 01:20:01 ; Search time 528.737 Seconds
(without alignments)
2304.231 Million cell updates/sec

Title: US-09-873-409-7

Perfect score: 2739

Sequence: 1 MILGILASLVNGACPLMPL.....SESKSAVOALEKDTPTRYSF 541

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2195239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -DEV=xlh

-MODEL=frame_p2n.model -DEV=xlh
-Q/cn2_1/USFT0.spool/US09873409/runat_27032003_115416_19216/app_query.fasta_1.7544
-DB=N Geneseq 101002 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -DOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09873409 @CGN 1 1 4247 @runat_27032003_115416_19216 -NCPU=6 -ICPU=3
-NO_XUPXY -NO_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=10 -XGAPEXT=0.5 -FGAPOPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 101002.*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2437	89.0	3699	24	ABK83223	Human transporter
2	1566.5	57.2	4189	21	AAZ49334	Murine multidrug r
3	1566.5	57.2	4189	24	ABA94367	Mouse BCRP DNA rel
4	1566.5	57.2	4313	14	AAQ38950	Mouse multidrug re
5	1561.5	57.0	4279	22	AAQ38950	Dog P-glycoprotein
6	1557.5	56.9	4279	22	AAQ38950	Dog P-glycoprotein
7	1557.5	56.9	4279	22	AAQ38950	Dog P-glycoprotein
8	1557	56.8	4195	22	AAF86128	Cynomolgous monke
9	1555.5	56.8	3860	21	AAZ49332	Human wild-type mu
10	1555.5	56.8	3860	24	ABA94365	Human BCRP DNA rel
11	1552.5	56.7	4233	21	AAZ90198	Rat mdrlb2 (multis
12	1552.5	56.7	4233	22	AAF27498	Rat mdrlb2 multidr
13	1552.5	56.7	4279	22	AAQ34888	Dog P-glycoprotein
14	1552.5	56.7	4317	22	AAQ34889	Dog P-glycoprotein
15	1552.5	56.7	4669	19	AAV32645	Human P glycoprote
16	1552.5	56.7	4669	24	ABK52041	cDNA encoding huma
17	1548.5	56.5	4186	22	AAF86127	Cynomolgous monke
18	1548	56.5	4254	24	ABK63517	Rat sequence diffe
19	1546.5	56.5	3860	21	AAZ49333	Human G185V mutant
20	1546.5	56.5	3860	24	ABA94366	Human BCRP DNA rel
21	1546.5	56.5	4349	22	AAH57442	Human intestine ce
22	1546.5	56.5	4646	15	AAQ72872	Human multidrug re
23	1546.5	56.5	4646	21	AAZ94738	Human ATP binding
24	1546.5	56.5	4646	24	AAZ38994	Human mdrl gene.
25	1546.5	56.5	4646	24	ABL68592	Kidney cancer rela
26	1546.5	56.5	4646	24	ABL68880	Kidney cancer rela
27	1546.5	56.5	6505	17	AAZ13394	Hybrid vector pSF-
28	1546.5	56.5	8630	21	AAZ24041	Retroviral M4 mdr-
29	1546.5	56.5	8630	21	AAZ24042	Retroviral vector
30	1545	56.4	4264	19	AAV66533	Mutated human P-gl
31	1545	56.4	4264	19	AAV66533	Mutated human P-gl
32	1544.5	56.4	4669	8	AAQ70522	Sequence of human
33	1543.5	56.4	4378	11	AAQ04522	Multidrug Resistan
34	1543	56.3	4425	21	AAZ52048	Rat multidrug resi
35	1541	56.3	4369	21	AAZ52047	Rat multidrug resi
36	1540.5	56.2	4669	14	AAQ52726	Sequence of human
37	1539	56.2	3840	24	ABL91687	Human polynucleoti
38	1539	56.2	3988	21	AAZ88973	Human MDR-1 DNA.
39	1536	56.1	4788	24	AAZ49335	Murine multidrug r
40	1536	56.1	4788	24	ABA94368	Mouse BCRP DNA rel
41	1455.5	53.1	3924	21	AAZ94742	Human ATP binding
42	1455.5	53.1	3924	21	AAZ88974	Human MDR-3 DNA.
43	1455.5	53.1	3924	24	ABN95801	Gene #2299 used to
44	1421.5	51.9	3912	24	ABK63653	Rat sequence diffe
45	1350	45.6	4776	21	AAZ94744	Human ATP binding

ALIGNMENTS

RESULT 1

ABK83223

ID ABK83223 standard; cDNA; 3699 BP.

XX ABK83223;

AC ABK83223;

XX 27-AUG-2002 (first entry)

XX Human transporter and ion channel, TRICH14, Incyte ID 7472030CB1, cDNA.
XX Human; ss; gene; transporter and ion channel; TRICH; transport disorder;
XX neurological disorder; muscle disorder; immunological disorder; cancer;
XX scleroderma; systemic lupus erythematosus; allergy; leukaemia;
XX cell proliferative disorder; cervical cancer; breast cancer;
XX neurodegenerative disorder; Parkinson's disease; Alzheimer's disease;
XX myotonic dystrophy; catatonias; endocrine disorder; diabetes;
XX Grave's disease; gastrointestinal disorder; Crohn's disease;
XX renal disorder; Good pasture's syndrome; viral infection; cirrhosis;

KW bacterial infection; fungal infection; parasitic infection;
 KW protozoal infection; helminthic infection; cardiovascular disorder;
 KW atherosclerosis; hepatic disease.

XX Homo sapiens.

XX WO200240541-A2.

XX 23-MAY-2002.

XX 25-OCT-2001; 2001WO-US46055.

XX 27-OCT-2000; 2000US-243989P.

XX 03-NOV-2000; 2000US-245904P.

XX 09-NOV-2000; 2000US-247673P.

XX 17-NOV-2000; 2000US-249661P.

XX 20-NOV-2000; 2000US-252232P.

XX 01-DEC-2000; 2000US-250790P.

XX (INCV-) INCVTE GENOMICS INC.

XX Tang YT, Yue H, Nguyen DB, Hafalia AJA, Elliott VS, Lu Y;

XX Wallia NK, Yao MG, Baughn MR, Gandhi AR, Ding L, Sanjanwala M;

XX Ramkumar J, Arvizu C, Gietzen KJ, Lal PG, Azimzai Y, Khan FA;

XX Thangavelu K, Thornton M, Lu DAM, Tribouley CM, Warren BA;

XX Ison CH, Das D, Raumann BE, Policky JL, Kearney L;

XX WPI; 2002-463570/49.

XX P-PSDB; ABG61544.

XX New transporters and ion channels (TRICH) polypeptides, useful for
 PT diagnosing, preventing, and treating disorders associated with an
 PT abnormal expression or activity of TRICH, e.g. immunological, muscular
 PT or renal disorders

XX Claim 5; Page 172; 178pp; English.

XX The invention relates to human transporters and ion channels (TRICH)
 CC polypeptides, a naturally occurring amino acid sequence 90 % identical to
 CC TRICH, a biologically active fragment of TRICH or an immunogenic fragment
 CC of TRICH. Also included are an isolated polynucleotide encoding TRICH,
 CC a recombinant polynucleotide comprising a promoter sequence operably
 CC linked to the TRICH polynucleotide, a cell transformed with the
 CC recombinant polynucleotide, a transgenic organism comprising the
 CC recombinant polynucleotide, an isolated antibody that binds specifically
 CC to TRICH, and screening for compounds which bind to TRICH, modulate
 CC TRICH, modulate TRICH expression or are ant/agonists of TRICH.
 CC The polypeptides are useful for diagnosing, treating, and
 CC preventing transport, neurological, muscle, immunological disorders
 CC (e.g. scleroderma, systemic lupus erythematosus, allergies), cell
 CC proliferative disorders such as cancers (e.g. leukaemia, cervical or
 CC breast cancers), neurodegenerative disorders (e.g. Parkinson's disease,
 CC Alzheimer's disease), muscular disorders (e.g. myotonic dystrophy,
 CC catatonka), endocrine disorders (e.g. diabetes, Grave's disease),
 CC gastrointestinal disorders (e.g. Crohn's disease), renal disorders
 CC (e.g. Good pasture's syndrome), vital, bacterial, fungal, parasitic,
 CC protozoal and helminthic infections, cardiovascular disorders (e.g.
 CC atherosclerosis), or hepatic diseases (e.g. cirrhosis) and many
 CC other diseases and disorders detailed in the specification. They can also
 CC be used in assessing the effects of exogenous compounds on the
 CC expression of nucleic acid and amino acid sequences of transporters and
 CC ion channels. TRICH or its fragments may also be used in screening for
 CC compounds that specifically bind to and modulate the activity of TRICH.
 CC The polynucleotides can be used to create knock-in humanised animals or
 CC transgenic animals to model human disease. The present sequence
 CC encodes a TRICH protein.

XX Sequence 3699 BP; 1116 A; 707 C; 860 G; 1016 T; 0 other;

Alignment Scores:

Fred. No.: 1.45e-241 Length: 3699
 Score: 2437.00 Matches: 495
 Percent Similarity: 90.71% Conservative: 3

Best Local Similarity: 90.16% Mismatches: 9
 Query Match: 88.97% Indels: 42
 DB: 24 Gaps: 3

US-09-873-409-7 (1-541) x ABR83223 (1-3699)

QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
 DB 145 ATGATCTCTGGGTATATCTGACATCATCTGTTCAATGGAGCCTGCCTTCTTAATGCCACTG 204
 QY 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAenThrTyr 40
 DB 205 TGTATAGGAGAAATAGTATACCTTATTAGTGGATGCTAGTCCACACACAACA 264
 QY 41 SerPhePheArg-----LeuThr 46
 DB 265 AATTATCAGAACTGTACTCAGTCTCAAGAGAGCTGAATGAAGATATGACTCTGTTGACC 324
 QY 47 LeuTyrTyrValGlyIleGlyValAlaLeuIlePheGlyTyrIleGlnIleSerLeu 66
 DB 325 CTGTATTATGTTGGATAGTGTGCTGCTTGTATTTTGGTTACATACAGATTTCCTTG 384
 QY 67 TrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGlnPhePheHisSerVal 86
 DB 385 TGGATTATTAACCTGCAGCAGCAGACCAAGAGGATTTCGAAACAGTTTTTTCATTAGTT 444
 QY 87 LeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyLeuLeuAsnThrArgMet 106
 DB 445 TTGGCAGACGACATCGCTGGTTTGTAGTGTGATCGTGGTGAACCTTAACATCTGCATG 504
 QY 107 Thr---AspIleAspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGln 125
 DB 505 ACAGATGACATTGACAAATCAGTGTGTTGGAGATAAGATTCTCTGTTGTTTCAA 564
 QY 126 AsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLysGlyTrpLysLeuThr 145
 DB 565 AACATGCTACTATTTTCGATTGGCTGGCAGTTGGTGGTGAAGGGCTGGAAACCTCACC 624
 QY 146 LeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAlaAlaCysSerArgMet 165
 DB 625 CTAGTCACTCTATCCAGCTCTCTCTTATATGCTTCAGCGCAGCATGTTCTAGATG 684
 QY 166 ValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGlu 185
 DB 685 GTCATCTCTATGACCACTAAGGAATTAAGTGCCTATTCAAAGCTGGGCTGGGAGAA 744
 QY 186 GluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGln 205
 DB 745 GAAGTCTTGTCAATCCGACAGTCATAGCCTTTAGGGCCCGAGGAGAAAGAACTTCAA 804
 QY 206 ArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeu 225
 DB 804 ----- 804
 QY 226 SerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 245
 DB 805 -----AGGTATACACAGAACTCAAGATGCAAGAGATTTCG 843
 QY 246 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGly 265
 DB 844 ATAAAAAGGACTATAGCTCAAAAGTGTCTCTTGGTGGCTGTGTACTTCTTTATGAATGA 903
 QY 266 ThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 285
 DB 904 ACCTATGGACTTGTCTTTTGGTATGGAACCTCTTGATTTCTTATNGAGAACCTGGATAT 963
 QY 286 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 305
 DB 964 ACCATCGGACTGTCTCTGCTGTTTCTTTAGTGTAAATCCATAGTAGTTATTTCATTGGA 1023
 QY 306 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 325
 DB 1024 GCAGCAGTCCCTCACCTTTGAAACCTTCGCAATAGCCCGAGGAGCTGCTCTTTCATATTTTC 1083

QY 326 GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 345
 DB 1084 CAGGTTATTGATAAGAAACCCAGTATAGGTAACTTTTCACAGCTGGATATAACCTGAA 1143
 QY 346 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 365
 DB 1144 TCATAGAAGGAAGTGTGAATTTAAATGTTCTTTCAATTATCCATCAAGGCCATCT 1203
 QY 366 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 385
 DB 1204 ATCAAGATTCGAAAGGTCTGAATCTCGGAATTAAGTCTGGAGACAGCTGGCTTGGTC 1263
 QY 386 GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro 405
 DB 1264 GGTCTCAATGGCAGTGGGAAGAGTAGCTCAGCTTCTGCAGAGGTTATATGATCCG 1323
 QY 406 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 425
 DB 1324 GATGATGGCTTTATCATGTGGTGAATGAGATGACATCAGAGCTTTAAATGTGGCGCATTA 1383
 QY 426 ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn 445
 DB 1384 CGAGACCATATTGGATGGTGTAGTCAAGAGCTGTTTGTTCGGGACCACCATCACTAAC 1443
 QY 446 AsnIleLysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGlu 465
 DB 1444 AATATCAAGTATGGACGAGATGATGTGACTGTAGTGAAGAGATGAGAGAGCAGCAAGGAA 1503
 QY 466 AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 485
 DB 1504 GCAAAATCGCTATGATTTTATCATGGAGTTCCTTAATAAATTTAATACATTTGGTAGGGAA 1563
 QY 486 LysGlyValGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal 505
 DB 1564 AAGGAGCTCAAAATGATGGAGGCGAGAACAGAGATCGCAATTCCTCTGCTCTAGTT 1623
 QY 506 ArgAsnProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 525
 DB 1624 CGAAACCCCAAGATTCGTATTTAGATGAGGTACGTCTGCCCTGGATTCCAGAAAGCAAG 1683
 QY 526 SerAlaValGlnAlaLeuGluLys 534
 DB 1684 TCAGCTGTTTCAGCTGCACTGGAGAAG 1710
 RESULT 2
 AAZ49334
 ID AAZ49334 standard; cDNA; 4189 BP.
 AC AAZ49334;
 XX
 XX 14-MAR-2000 (first entry)
 DT
 XX Murine multidrug resistance-1 (MDR-1) cDNA.
 DE
 XX Multidrug resistance; MDR-1; P-glycoprotein;
 KW transmembrane efflux pump; haematopoietic stem cell; transduction;
 KW bone marrow transplantation; chemotherapy; radiation therapy; cancer;
 KW gene therapy; gene replacement; genetic defect; thalassaemia;
 KW Gauchier's disease; sickle cell anaemia; leukaemia; ex vivo expansion;
 KW cytokine; ds.
 XX
 XX Mus sp.
 XX
 XX Key Location/Qualifiers
 FH 1..3831
 CDS /tag= a
 FT /product= "Murine MDR-1 protein"
 FT
 FT
 FT
 XX
 PN W09961589-A2.
 XX
 XX 02-DEC-1999.
 PD
 XX

PF 27-MAY-1999; 99WO-US11825.
 XX
 PR 28-MAY-1998; 98US-0086988.
 XX
 PA (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
 XX
 XI Sorrentino B, Bunting K;
 PI
 XX WPI; 2000-072615/06.
 DR P-PSDB; AAY58188.
 XX
 XX Ex vivo expansion of hematopoietic stem cells transduced with a
 PT sequence encoding human multidrug resistance-1, used for bone marrow
 PT transplantation -
 XX
 PS Disclosure; Page 90-93; 113pp; English.
 XX
 CC This sequence represents cDNA encoding murine multidrug
 CC resistance protein MDR-1. MDR-1 is a transmembrane
 CC efflux pump, responsible for the export of drugs from cells,
 CC particularly cancer cells. The invention relates to transducing
 CC haematopoietic stem cells with nucleic acid encoding an MDR protein
 CC and culturing the modified cells. The modified haematopoietic stem
 CC cells are useful in bone marrow transplantation (to reconstitute
 CC haematopoietic systems in patients who have undergone chemotherapy or
 CC radiation therapy) and in ex vivo gene therapy of genetic defects in
 CC cells derived from haematopoietic stem cells, e.g., thalassaemia,
 CC Gauchier's disease, sickle cell anaemia or leukaemia. The modified
 CC cells can also be used to identify factors involved in regulating
 CC proliferation and differentiation in haematopoietic stem cells.
 CC Haematopoietic stem cells that express MDR-1 will be protected against
 CC chemotherapeutic agents, so can be engrafted while the patient is
 CC undergoing chemotherapy. Expansion of (rare) haematopoietic stem cells
 CC provides sufficient cells to permit standard biochemical analysis.
 CC Overexpression of MDR-1 allows cytokine-driven expansion of
 CC haematopoietic stem cells by at least 10-fold compared with a maximum
 CC of 4-fold in known procedures.
 XX
 SQ Sequence 4189 BP; 1204 A; 875 C; 1028 G; 1082 T; 0 other;
 Alignment Scores:
 Pred. No.: 2,06e-151 Length: 4189
 Score: 1566.50 Matches: 307
 Percent Similarity: 72.63% Conservative: 99
 Best Local Similarity: 54.92% Mismatches: 94
 Query Match: 57.19% Indels: 59
 DB: 21 Gaps: 6
 US-09-873-409-7 (1-541) x AAZ49334 (1-4189)
 QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyValaCysLeuProLeuMetProLeu 20
 DB 148 ATGATTCGGGAACCTCGCTGCTATTATCCATGGACATTACTTCCCTCTTGATGCTG 207
 QY 21 ValLeuGlyGluMetSerAspAsnLeuIle----- 30
 DB 208 GTGTTTGGAAACATGACAGATAGTTTACAAAAGCAGAGCCAGTATTCTGCCAAGCAT 267
 QY 31 -----SerGlyCysLeuValGlnThrAsnThr----- 39
 DB 268 ACTAATCAAAGTGGACCCCAACAGTAGTACTCTGATCATCAGCAACAGCAGTCTGGAGGAGAG 327
 QY 40 -----TyrSerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValaAla 56
 DB 328 ATGGCCATATAGCCCTAC-----TATTACCGGGATTGCTGCTGTG 372
 QY 57 LeuIlePheGlyTyrIleGlnIleSerLeuTyrIleIleThrAlaAlaArgGlnThrLys 76
 DB 373 CTCATAGTTGCTTACATCCAGGTTTCACCTTGTGCTGCTGAGCTGGAAGACAGATACAC 432
 QY 77 ArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSer 96
 DB 433 AAGATTAGGCAGAAAGTTTTTTCATGCTATATAATCAATCAGGAGATAGGCTGGTTGATGTG 492

Qy	97	CyeAspIleGlyLeuAsnThrArgMetThr	---AspIleAspLySIIleSerAspGly	115
Db	493	CATGATGTTGGGAGCTCAACCCGGCTCACAGATGATGCTCCAAAATTAATGACGGA	552	
Qy	116	IleGlyAspLySIIleAlaLeuLeuPheGlnAenMetSerThrPheSerIleGlyLeuAla	135	
Db	553	ATTGGTGACAAAATTGGGATGTTTTTTCAGTCCATAAACACCATTTTTAGCCGGTTTTATC	612	
Qy	136	ValGlyLeuVallySgLyTrpLySLeuThrLeuValThrLeuSerThrSerProLeuIle	155	
Db	613	ATAGGATTATATAAGTGGTGGAGCTAACCCCTTGTCATTTTGGCTGTCCAGCCCTCTTATT	672	
Qy	156	MetAlaSerAlaAlaAlaCysSerArgMetValIIleSerLeuThrSerLySLeuLeuSer	175	
Db	673	GGATTGTCATCTGCTTGTGGGCAAGGATTGACTTCATTACTAATAGGAACCTCCAG	732	
Qy	176	AlaTySerLySIIleAlaValAlaGluGluValLeuSerSerIleArgThrValIle	195	
Db	733	GCTTATGCAAAAGCTGGAGCAGTTGCTGAAGAAGTCTTAGCAGCATCAGAACTGTGATT	792	
Qy	196	AlaPheArgAlaGlnGluLySLeuGlnArgSerPheLeuLeuAsnIleThrArgTyr	215	
Db	793	GCCTTTGGAGGACAAACAGAAAGAACTTGAA	822	
Qy	216	AlaTrpPheTyrPheProGlnTrpLeuLeuSerCysValLeu	235	
Db	823	-----AGGTACAAT	831	
Qy	236	GlnAsnLeuLySIIleAlaLySIIleAspPheGlyIleLySIIleArgThrIleAlaSerLySIIleValSer	255	
Db	832	AAAAATTTAGAAAGCTAAATAATGTTGGCAATAAAGAAAGCTATCACAGCCAGCATTTGC	891	
Qy	256	LeuGlyAlaValTyrPhePheMetAenGlyThrTyrGlyLeuAlaPheTrpTyrGlyThr	275	
Db	892	ATAGGCATTCCCTACCTGTTGGTCTATGCATCATATGACCTGGCATTCCTGCTATGGGACA	951	
Qy	276	SerIleIleLeuAenGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePhe	295	
Db	952	TCCCTGTGCTCTCAATGAA-----TATCTATGGAGAGTCTTACTGTCTCTCTTC	1005	
Qy	296	SerValIleHisSerSeryCysIleGlyAlaAlaValProHisPheGluThrPheAla	315	
Db	1006	TCATTTTGTGGGACTTTAGTATTGCACACTTGGCCCAACATAGAACCTTTTGCA	1065	
Qy	316	IleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLySIIleProSerIleAsp	335	
Db	1066	AACGCACAGGGGCGAGCTTTGAAATCTTCAAGATAATTGATACGAGCCACCAAGCATTTGAC	1125	
Qy	336	AsnPheSerThrAlaGlyTrpLySProGluSerIleGluGlyThrValGluPheLySAsn	355	
Db	1126	AGCTTCTCAACAAAGGGCTACAAACAGACAGTATAATGGGAAACTTAGAGTTTTAAAAAT	1185	
Qy	356	ValSerPheAsnTyrProSerArgProSerIleLySIIleLeuLySIIleGlyValLeuAsnLeuArg	375	
Db	1186	GTTCACCTTCAACTACCATCGAAGCGAAGTTGAGATCTTGAAGGGCTCAATCTGAAG	1245	
Qy	376	IleLySIIleSerGlyGlyThrValAlaLeuValGlyLeuAenGlySerGlyLySIIleSerThrVal	395	
Db	1246	GTGAAGACGGACAGACGGTGGCTTTGGTTGGCAACAGTGGCTGTGGGAAAAAGCACAACT	1305	
Qy	396	ValGlnLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsn	415	
Db	1306	GTCACGCTGATGACAGAGGCTCTACGACCCCTCGGAGGGCGTGTGATGATCGACGGACAA	1365	
Qy	416	AspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGlu	435	
Db	1366	GACATCAGAACCATCAATGTGAGGTATCTGAGGGAGATCATTTGGTGTGGTGAATCAGAA	1425	
Qy	436	ProValLeuPheGlyThrThrIleSerAenAsnIleLySIIleTyrGlyArgAspValThr	455	
Db	1426	CCTGTGCTGTTGCCACCATCGCCGAGAACATTCGTATGCGCCGAGAGATGTCAACC	1485	

Qy	456	AspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPhe	475
Db	1486	ATGGATCAGATTTCAGAAAGCTGTCAAGGAAGCAATGCTCTATGACTTCATCATGAAC	1545
Qy	476	ProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLys	495
Db	1546	CCCCACCAATTTGACACCCCTGGTGTGTGAGAGAGGGGGCGACGCTGAGTGGGGGACAGAAA	1605
Qy	496	GlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGlu	515
Db	1606	CAGAGATCCCAATTCGCCGGGCCCTGGTCCGCAATCCAGAGATCTTTTGTGGACGAG	1665
Qy	516	AlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys	534
Db	1666	GCCACCTCAGCCCTGGATACAGAAAGTGAAGCTGTGTTGTCAGGCCGCACTGGATAAG	1722
RESULT	3		
ABA94367			
ID	ABA94367	standard; DNA; 4189 BP.	
XX	AC	ABA94367;	
XX	DT	26-MAR-2002 (first entry)	
XX	DE	Mouse BCRP DNA related seq Id No. 5.	
XX	KW	Stem cell; ATP transport protein; ATP-binding cassette; antiparkinsoni	
XX	KW	hepatotropic; neurodegenerative; cytotostatic; antianemic; muscular; BCRP	
XX	KW	cardiant; gene therapy; ds.	
XX	OS	Mus musculus.	
XX	Key	Location/Qualifiers	
XX	FT	1..3831	
XX	FT	/*tag= a	
XX	FN	WC200192877-A2.	
XX	PD	06-DEC-2001.	
XX	PF	30-MAY-2001; 2001WO-US17459.	
XX	PR	31-MAY-2000; 2000US-0584586.	
XX	PR	29-MAY-2001; 2001US-0866866.	
XX	PA	(SUUD-) ST JUDE CHILDREN'S RES HOSPITAL.	
XX	PI	Sorrentino B, Schuetz J;	
XX	DR	WPI; 2002-114368/15.	
XX	DR	P-PSDB; ABB07268.	
XX	PT	Identifying a stem cell, for treating e.g., muscular dystrophy,	
XX	PT	myocardial infarction, Parkinson's disease, or neurodegenerative	
XX	PT	disorders, comprises detecting the expression of an ATP transport	
XX	PT	protein (BCRP) by a cell	
XX	PS	Disclosure; Page 64-66; 87pp; English.	
XX	CC	The invention provides a method of identifying and/or isolating a stem	
XX	CC	cell that involves detecting the expression of an ATP transport protei	
XX	CC	containing a conserved ATP-binding cassette (BCRP) by a cell in a samp	
XX	CC	comprising stem cells. The isolated stem cells may be used in the	
XX	CC	treatment of diseases such as muscular dystrophy, degenerative liver	
XX	CC	disorder, myocardial infarction, Parkinson's disease, degenerative	
XX	CC	disorders of the brain, and for tissue regeneration or replacement.	
XX	CC	Haematopoietic cells can be used in bone marrow transplants (e.g., for	
XX	CC	treatment of leukemia) and for ex vivo gene therapy for treating blood	
XX	CC	diseases such as sickle cell anemia and thalassemia. The stem cells ca	
XX	CC	also be used as cell targets in gene therapy protocols. The present	
XX	CC	sequence represents a sequence related to the BCRP for which no releva	
XX	CC	information has been provided in the specification.	


```
XX 15-JUL-1986; 86US-0885951.
XX
XX 15-JUL-1986; 86US-0885951.
XX 06-FEB-1991; 91US-0652311.
XX (MASI ) MASSACHUSETTS INST TECHNOLOGY.
XX
XX Croop JM, Gros P, Housman DE;
XX
XX WPI; 1993-126077/15.
XX P-PSDB; AAR35199.
XX
XX DNA sequence which confers multi-drug resistance on sensitive
XX mammalian cells - used to preserve bone marrow cells during
XX chemotherapy to prevent infection
XX
XX Claim 1; Fig 8; 22pp; English.
XX
XX A cDNA library was constructed from mRNA which had been isolated
XX from a drug-sensitive mouse cell-line. Two mouse cDNA molecules
XX complementary to the mRNA species encoded by 2 related but distinct
XX mdr genes were isolated and cloned. One of the cDNA clones (lambda
XX Dril1) is a full-length cDNA clone for one member of the mdr gene
XX family. When incorporated into prokaryotic expression vector pDREX4
XX (which allows high levels of transcription of the cDNA when
XX introduced into mammalian cells), the clone was shown to confer the
XX multidrug resistance phenotype upon transfection into drug-sensitive
XX mammalian cells.
XX
XX SQ Sequence 4313 BP; 1241 A; 904 C; 1058 G; 1110 T; 0 other;

Alignment Scores:
Pred. No.: 2.15e-151 Length: 4313
Score: 1586.50 Matches: 307
Percent Similarity: 72.63% Conservative: 99
Best Local Similarity: 54.92% Mismatches: 94
Query Match: 57.19% Indels: 59
DB: 14 Gaps: 6

US-09-873-409-7 (1-541) x AAQ38950 (1-4313)
Qy 1 MetileLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 257 ATGATTCTGGGAACCTCGCTGCTATTATCCATGGAACATTACTTCCCTCTTGATCGTG 316
Qy 21 ValLeuGlyGluMetSerAspAsnLeuLeu----- 30
Db 317 GTGTTTGGMAACATGACAGATAGTTTTACAAAAGCAGAGCCAGTATTCTGCCAAGCATT 376
Qy 31 -----SerGlyCysLeuValGlnThrAsnThr----- 39
Db 377 ACTAATCAAAAGTGACCCCAACAGACTCTGATCATCATGACACAGCAGCTGGAGGAAGAG 436
Qy 40 -----TyrSerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAla 56
Db 437 ATGCCCATATACGCTAG-----TATTACACCGGGATGGTGGTGGTG 481
Qy 57 LeuIlePheGlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLys 76
Db 482 CTCATAGTTGGCTACATCCAGGTTTCACTTGTGTGGCTGGCAGCTGGAAGACAGATACAC 541
Qy 77 ArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSer 96
Db 542 AAGATTAGGCGAAGATTTTCCATGCTATATGAATCAGGAGATAGCTGGTTGATGTG 601
Qy 97 CysAspIleGlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGly 115
Db 602 CATGATGTTGGGAGGAGCTCAACACCGGCTCACAGATGATGCTCCAAATTAATGACGGA 661
Qy 116 IleGlyAspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAla 135
Db 662 ATTGCTGACAAAATTTGGGATGTTTTTTTCAGTCCATAACCAACCATTTTATAGCGGTTTATC 721
```

```
Qy 136 ValGlyLeuValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIle 155
Db 722 ATAGGATTATTAAGTGGTTGGAGGCTAACCCCTTGTCATTTGGCTGTCAGCCCTTATT 781
Qy 156 MetAlaSerAlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSer 175
Db 782 GGATTTCATCTGCTTGTGGGCAAAAGGATTGACTTCATTACTAATAAGGAATCCAG 841
Qy 176 AlaTyrSerLysAlaGlyAlaValAlaGluValLeuSerSerIleArgThrValIle 195
Db 842 GCTTATGCAAAAGCTGGAGCAGTTGCTGAAGAGTCTTAGCAGCCATCAAGACTGGATT 901
Qy 196 AlaPheArgAlaGlnGlnLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyr 215
Db 902 GCCTTTGGAGGACACAGAGGAACTTGAA----- 931
Qy 216 AlaTrpPheTyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThr 235
Db 932 -----AGGTACAAT 940
Qy 236 GlnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSer 255
Db 941 AAAAATTTAGAGAGAGCTAAATAATGTTGCATAAAGAAAGCTATCACAGCCAGCATTCG 1000
Qy 256 LeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThr 275
Db 1001 ATAGGCAATTCCTGCTGCTGTTGCTATGATCATATGCACTGGCATTCGGTATGGGACA 1060
Qy 276 SerLeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePhe 295
Db 1061 TCCTTGGTCTCTCAAAATGAA-----TATTCTATTGGAGAGTGTCTACTGCTCTCTTC 1114
Qy 296 SerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAla 315
Db 1115 TCTATTTTGTGGGACTTTTATGATTGGACACTTGGCCCCCAACACATAGAAGCTTTGCA 1174
Qy 316 IleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAsp 335
Db 1175 AACGCACAGGGGCGCCCTTTGAAATCTTCAAGATTAATTAACAGCCACGATTCGATGAC 1234
Qy 336 AsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsn 355
Db 1235 AGCTTCTCAACAAAGGGCTACAAACACAGACAGTATATATGGAAACTTAGAGTTAAAAAT 1294
Qy 356 ValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArg 375
Db 1295 GTTCACCTTCAACTTACCCTACCATCGAGAGCGAAGTTCAGATCTTGAAGGGCCTCAATCTGAAG 1354
Qy 376 IleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrVal 395
Db 1355 GTGAAGAGCGGACAGACGGTGGCTTGGTTGGCAACAGTGGCTGTGGNAANAACCAACT 1414
Qy 396 ValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsn 415
Db 1415 GTCCAGCTGATGCAGAGGCTTACGACCCCTCGAGGGGCTGTGTATCGACGGACAA 1474
Qy 416 AspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGlu 435
Db 1475 GACATCAGAACCAATCAATATGAGGTATCTCGAGGAGATCATTTGGTGTGGTACGAGAA 1534
Qy 436 ProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspValThr 455
Db 1535 CTTGTGCTGTTGCCACACGATCGCCGAGACATTCGTTATGCGCCGAGAGATGTCACC 1594
Qy 456 AspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPhe 475
Db 1595 ATGCATGAGATTGAGAAAGCTGTCAAGGAAGCAATGCTATGCTTATCATGAACATG 1654
Qy 476 ProAsnLysPheAsnThrLeuValGlyGluLysGlyValAlaGlnMetSerGlyGlnLys 495
Db 1655 CCCACCAATTTGACACCTGGTTGGTGAGAGAGGGCGCAGCTGAGTGGGGGACAGAAA 1714
```


Db 926 GCTTCTTATTGATCTATGTCATCATATGCTCTGGCTTTCTGGTATGGACCTCTCTGGTC 985
Qy 279 LeuAsnGlyGluProGlyThrValLeuAlaValPheSerValIle 298
Db 986 CTCTCCAGTGAA-----TATACATATTGGACAGTACTCACTGCTCTCTCTCTCTGATTA 1039
Qy 299 HisSerSerTyrcysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArg 318
Db 1040 ATTTGGGCTTTAGTATTGGACAGCATCCCAAGCATTTGAAGCATTTGCAACGCAAGA 1099
Qy 319 GlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSer 338
Db 1100 GGAGCAGCTTATGAATCTTCAAGATATTGACAAATAAACAAGCATTTGACAGCTATTTCG 1159
Qy 339 ThrAlaGlyThrLysProGluSerIleGlyThrValGluPheLysAsnValSerPhe 358
Db 1160 AAGAGTGGACATAAACCATGATATTAAAGGAAATTTGGAATTTCAAAAATGTTCACTTC 1219
Qy 359 AsnTyrrProSerArgProSerIleLysIleLysGlyLeuAsnLeuArgIleLysSer 378
Db 1220 AGTTACCTCTTCGAAAGAGATTAAAGATCTTAAAGGCTCTCAACCTGAAGGTTTCAGAGT 1279
Qy 379 GlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValGlnLeu 398
Db 1280 GGGCAGACAGTGGCGCTGGTTGGGAACAGTGGCTGCGGGAAGACACGACCGCTGCAGCTG 1339
Qy 399 LeuGlnArgLeuTyrrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArg 418
Db 1340 ATGCAGAGGCTCTATGACCCACAGATGGCATGCTGTGTATTGATGACAGGACATTAGG 1399
Qy 419 AlaLeuAsnValArgHisTyrrArgAspHisIleGlyValValSerGlnGluProValLeu 438
Db 1400 ACCATAAATGTAAGGCATCTTCGGGAATTAATCTGTTGTGTGAGTCAGGAGCTGTGTTG 1459
Qy 439 PheGlyThrThrIleSerAsnAsnIleLysTyrrClyArgAspAspValThrAspGluGlu 458
Db 1460 TTTGCCACACGATAGTGAATAACATTCGCTATGCGCGCAAAATGTCACCATGGATGAG 1519
Qy 459 MetGluArgAlaAlaArgGluAlaAsnAlaTyrrAspPheIleMetGluPheProAsnLys 478
Db 1520 ATTGAGAAAGCTGTAAAGGAAGCCAAATGCCATATATTATCATGAACACTACCTAATAAA 1579
Qy 479 PheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIle 498
Db 1580 TTTGACACTCTGTTGGAGAGAGAGGCGCCAGCTGAGTGGTGACAGAAACAGAGATC 1639
Qy 499 AlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSer 518
Db 1640 GCCATTGCTCGGGCCCTGGTTGCGCAACCCCAAGATTCTTCTGCTGGATGAGCAACGTCA 1699
Qy 519 AlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
Db 1700 GCTCTGGACACTGAAGTGAAGCAGTGGTTGAGTGGCCCTGGATAAG 1747

RESULT 6

AAD03504

ID AAD03504 standard; cDNA; 4279 BP.

XX AAD03504;

XX AAD03504;

XX 13-JUN-2001 (first entry)

XX Dog P-glycoprotein (PGP) allelic variant (Genotype A) cDNA.

XX Dog; P-glycoprotein allelic variant; multidrug transporter; MDR1;
KW drug bioavailability; transgenic animal; genetic model; ss.

XX Canis familiaris.

XX Key Location/Qualifiers

XX CDS 17..3862

XX /*tag= a

XX /product= "Dog P-glycoprotein (PGP) allelic variant

allele (Genotype A) protein" replace (607, C) /*tag= b
WO200123540-A2.
XX 05-APR-2001.
XX 28-SEP-2000; 2000WO-US26767.
XX 28-SEP-1999; 99US-0156510.
XX (GENT-) GENTEST CORP.
XX Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;
XX WPI; 2001-235373/24.
XX P-PSDB; AAE00308.
XX New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
XX for determining the bioavailability of drugs and for screening for dog
XX PGP inhibitors -
XX Claim 9; Page 85-90; 11pp; English.
XX The invention relates to dog P-glycoprotein (PGP) also referred
XX as multidrug transporter (MDR1) and nucleic acids encoding them.
XX The invention also includes fragments and biologically functional
XX variants of dog P-glycoprotein. PGP and their nucleic acids are
XX useful for determining the bioavailability of drugs and for
XX screening PGP inhibitors. They are useful for the diagnosis and
XX treatment of conditions characterised by PGP activity, by
XX reducing or increasing PGP activity in a cell. PGP nucleic acids
XX are used as oligonucleotide probes. Complements of PGP nucleic
XX acids are useful as antisense oligonucleotides, to induce a PGP
XX 'knockout' phenotype. They are used to prepare a non-human
XX transgenic animal, which are valuable as genetic models for
XX human diseases.
XX The present sequence is dog P-glycoprotein (PGP) allelic variant
XX (Genotype A) cDNA. The PGP enzyme functions as an efflux pump
XX exporting small molecules across the cell membrane. This enzyme
XX is a member of the ABC transporter family.
XX Sequence 4279 BP; 1295 A; 833 C; 1008 G; 1143 T; 0 other;
SQ
Alignment Scores:
Pred. No.: 1,81e-150 Length: 4279
Score: 1557.50 Matches: 303
Percent Similarity: 72.12% Conservative: 98
Best Local Similarity: 54.50% Mismatches: 104
Query Match: 56.86% Indels: 51
DB: 22 Gaps: 5
US-09-873-409-7 (1-541) x AAD03504 (1-4279)
Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 167 AUGTTGGTGGGACAAATGGCTGCCATCATCCATGGAGCTGCATCCCTCTCATGATGCTG 226
Qy 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu----- 34
Db 227 GTTTTGGAAACATGACAGATAGCTTTGCAAAATGCAAGAAATTTCAAGAAACAAACTTTT 286
Qy 35 -----ValGlnThrAsnThrTyrrSerPhePhearg----- 44
Db 287 CCAGTTATAATTAATGAAGTATTACGAACAATACACAACTTTCATCAACCATCTGGAG 346
Qy 45 -----LeuThrLeuTyrrValGlyIleGlyValAlaAlaLeuIlePhe 59
Db 347 GAGAAATGACACAGCTATGCCTATTATTACAGTGGATCGGTCTGGCTGGTGGCT 406
Qy 60 GlyTyrrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg 79

Db 407 GCTTACATCCAGGTTTCATTCTGGTGCCTGGCAGCAGGAGACAGATACACTCAAAATTAGA 466
Qy 80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIle 99
Db 467 AAACAATTTTTCATGCTATCATCGCAGCAGGAGATTGGCTGGTTTGACGTGCATGACGTT 526
Qy 100 GlyLeuLeuAenThrArgMetThr---AspIleAspIleSerAspGlyIleGlyAsp 118
Db 527 GGGGAGCTTAACACCCGGCTCACACACATGCTCCAAATCAATGAAGGAATTGGCCGAC 586
Qy 119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
Db 587 AAATGGAATGTTCTTCAATCATAGCAACATTTTTCACCGGTTTATAGTGGGGTTT 646
Qy 139 ValLysGlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
Db 647 ACACGTGGTGGAGCTAAACCTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACATTCA 706
Qy 159 AlaAlaIaCysSerArgMetValIleSerLeuThrSerLysLeuLeuSerAlaTyrSer 178
Db 707 GCGCCATCTGGGCAAGATACTATCTTCATTACTGATAAAGAACTCTTGGCCCTATGCA 766
Qy 179 LysAlaGlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArg 198
Db 767 AAAGCTGGAGCAGTAGCTGAAGAGCTTTAGCAGCAATCAGAACTGTGATTCCTTTGGA 826
Qy 199 AlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTyrPhe 218
Db 827 GGACAAAGAAAGAACTTGAA----- 847
Qy 219 TyrPheProGlnTyrLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeu 238
Db 848 -----AGTCAACAACAAAATTATA 865
Qy 239 LysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAla 258
Db 866 GAAGAAGCTAAAGAAATTTGGGATAAAGAAAGCTACACGGCCAAACAACTTCTATTTGGTGCC 925
Qy 259 ValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIle 278
Db 926 GCTTCTTATTGATCATCATCATGCTCTGGCTTCTGGTATGGGACCTCTTGGTC 985
Qy 279 LeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPheSerValIle 298
Db 986 CTCCTCCAGTGAA-----TATTCTATTGGACAAGTACTCACTGTCTCTTCTTCTGATTATA 1039
Qy 299 HisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArg 318
Db 1040 ATTGGGGCTTTTAGTATTGGACAGGCATCCCCAAGCATTTGAAGCATTTTGAACACCAAGA 1099
Qy 319 GlyAlaAlaPheHisPheIlePheGlnValIleAspLysProSerIleAspAsnPheSer 338
Db 1100 GGAGCAGCTTATGAATCTTCAAGATATTGCAATTAACCAAGCATTTGACAGCTATTGCG 1159
Qy 339 ThrAlaGlyTyrLysProGluSerIleGlyThrValGluPheLysAsnValSerPhe 358
Db 1160 AAGAGTGACATAAACCAAGATAATATTAAAGGAAATTTGGAATTTCAAAATGTTCACTTC 1219
Qy 359 AsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSer 378
Db 1220 AGTTACCTCTCTCGAAAGAAAGTAAAGATCTTAAAGGGTCTCAACCTGAAGGTTTCAGAT 1279
Qy 379 GlyGluThrValAlaLeuValGlnGlySerGlyLysSerThrValValGlnLeu 398
Db 1280 GGGCAGACAGTGGCGTGGTGGGAACAGTGGCTGGGGAAGAGCAGCACCGCTGCAGCTG 1339
Qy 399 LeuGlnArgLeuTyrAspProAspGlyPheIleMetValAspGluAsnAspIleArg 418
Db 1340 ATGCAGAGGCTTATGACCCACAGATGGCATGGTCTGTATTGATGGACAGACATTAGG 1399
Qy 419 AlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeu 438
Db 1400 ACCATAAATGTAAAGCATCTTCGGGAAATTTACTGGTGTGGTGGAGCAGGAGCTGTGTTG 1459

Qy 439 PheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGlu 458
Db 1460 TTTGCCACCATAGCTGAAACATTCGCTATGCGCCGCAAAATGTCCCATGGATGAG 1519
Qy 459 MetGluArgAlaAlaAaGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLys 478
Db 1520 ATTGAGAAGCTGTTAAGGAAGCCATGCCCTATGATTTTATCATGAAACTACCTAATAA 1579
Qy 479 PheAsnThrLeuValGlyLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIle 498
Db 1580 TTTGACACTCTGGTTGGAGAGAGAGGGGCCAGCTGAGTGGTGGACAGAAACAGAGAATC 1639
Qy 499 AlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSer 518
Db 1640 GCCATTGCTCGGGCCCTGGTTTCGAACCCCAAGATTCTTCTGCTGATGAGGCAACGTC 1699
Qy 519 AlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
Db 1700 GCTCTGGACACTGAAAGTGAAGCAGTGGTTTCAGGTGGCCCTGGATAG 1747
RESULT 7
AAD03505
ID AAD03505 standard; cDNA; 4279 BP.
XX
AC AAD03505;
DT 13-JUN-2001 (first entry)
XX
DE Dog P-glycoprotein (PGP) allelic variant (Genotype B) cDNA.
XX
KW Dog; P-glycoprotein allelic variant; PGP; multidrug transporter;
KW MDRI; drug bioavailability; transgenic animal; genetic model; ss.
XX
OS Canis familiaris.
FH Key Location/Qualifiers
FT CDS 17..3862
FT /*tag= a
FT /product= "Dog P-glycoprotein (PGP) allelic variant
FT (Genotype B) protein"
FT replace (91, T)
FT /*tag= b
FT /*tag= c
FT /*tag= c
XX WO200123540-A2.
XX
XX 05-APR-2001.
XX
XX 28-SEP-2000; 2000WO-US26767.
XX
XX 28-SEP-1999; 99US-0156510.
XX
XX (GENT-) GENTEST CORP.
XX
XX Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;
XX
XX WPI; 2001-235373/24.
XX P-PSDB; AAE00309.
XX
XX New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
XX for determining the bioavailability of drugs and for screening for dog
XX PGP inhibitors -
XX
XX Claim 9; Page 93-99; l1lpp; English.
XX
XX The invention relates to dog P-glycoprotein (PGP) also referred
XX as multidrug transporter (MDRI) and nucleic acids encoding them.
XX The invention also includes fragments and biologically functional
XX variants of dog P-glycoprotein. PGP and their nucleic acids are
XX useful for determining the bioavailability of drugs and for
XX screening PGP inhibitors. They are useful for the diagnosis and

CC treatment of conditions characterised by PGP activity, by
 CC reducing or increasing PGP activity in a cell. PGP nucleic acids
 CC are used as oligonucleotide probes. Complements of PGP nucleic
 CC acids are useful as antisense oligonucleotides, to induce a PGP
 CC 'knockout' phenotype. They are used to prepare a non-human
 CC transgenic animal, which are valuable as genetic models for
 CC human diseases.
 CC The present sequence is dog P-glycoprotein (PGP) allelic variant
 CC (Genotype B) cDNA. The PGP enzyme functions as an efflux pump
 CC exporting small molecules across the cell membrane. This enzyme
 CC is a member of the ABC transporter family.

SX Sequence 4279 BP; 1296 A; 833 C; 1008 G; 1142 T; 0 other;

Alignment Scores:

Pred. No.: 1,81e-150 Length: 4279
 Score: 1557.50 Matches: 303
 Percent Similarity: 72.12% Conservative: 98
 Best Local Similarity: 54.50% Mismatches: 104
 Query Match: 56.86% Indels: 51
 DB: 22 Gaps: 5

US-09-873-409-7 (1-541) x AAD03505 (1-4279)

QY 1 MetIleuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
 DB 167 ATGTTGGTGGGACATGGCTGCATCATCCATGGAGTGCATCTCCTCATGATGCTG 226
 QY 21 ValLeuGlyLeuMetSerAspAsnLeuIleSerGlyCysLeu----- 34
 DB 227 GTTTTGGAAACATGACAGATAGCTTTGCAATGCGAGGAATTTCAAGAAACAAACTTTT 286
 QY 35 -----ValGlnThrAsnThrTyrSerPheArg----- 44
 DB 287 CCAGTTATAATTAATGAAGTATTACGAACAATACACAACTTCACCACTGGAG 346
 QY 45 -----LeuThrLeuTyrValGlyIleGlyValAlaLeuIlePhe 59
 DB 347 GAGGAATGACCAGCTATGCTATTATTACAGTGGGATCGTGGCTGGCTGGCT 406
 QY 60 GlyTyrIleGlnIleSerLeuThrIleThrAlaAlaArgGlnThrIysArgIleArg 79
 DB 407 GCTTATCATCCAGGTTTCATCTGCTGGCAGCAGGAGACAGATACTCAAAATTAGA 466
 QY 80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIle 99
 DB 467 AACAATTTTTTCATGCTATACGACAGAGATGGCTGGTTGACGTGCATGACGTT 526
 QY 100 GlyGluLeuAsnThrArgMetThr---AspIleAspIysIleSerAspGlyIleGlyAsp 118
 DB 527 GGGGAGCTTAACACCCGGCTCACAGACGATGCTCCAAATCAATGAGGAATGGCGAC 586
 QY 119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
 DB 587 AAAATTGGAATGTTCTTCAATCAATAGCAACATTTTACCCGTTTATAGTGGGGTTT 646
 QY 139 ValLysGlyTyrIysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
 DB 647 ACACGTGGTGGGAAGTAACCCCTGTGATTTGGCCATCAGCCCTGTTCTGGCCTTCA 706
 QY 159 AlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
 DB 707 GCGCCATCTGGCAAGATACCTATCTTCTTACTGATTAAGAACTCTTGGCCTATGCA 766
 QY 179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198
 DB 767 AAAGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGAATCTGTGATGCTTTGGA 826
 QY 199 AlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPhe 218
 DB 827 GGACAAAGAAAGAACTTGAA----- 847
 QY 219 TyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeu 238

DB 848 -----AGGTACACAAAAATTTA 865
 QY 239 LysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAla 258
 DB 866 GAAGAAGCTAAAGGAATTGGGATAAGAAAGTATCACGCCCAACATTTCTATGTGTGC 925
 QY 259 ValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIle 278
 DB 926 GCTTCTTATTGATCATCATCATATGCTCTGGCTTCTGGTATGGACCTCTCTGGTC 985
 QY 279 LeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIle 298
 DB 986 CTCTCCAGTGAA-----TATTCTATTGGCAAGTACTCACTGTCTCTTCTCTGATT 1039
 QY 299 HisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArg 318
 DB 1040 ATTGGGGCTTTTAGTATTGGACAGGCATCCCAAGCATTTGAAGCATTTGCCAACGGA 1099
 QY 319 GlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSer 338
 DB 1100 GGAGCAGCTTATGAATCTTCAAGATATTTGCAATAAACCAAGCATTTGACAGCTATTC 1159
 QY 339 ThrAlaGlyTyrLysProGluSerIleGlyThrValGluPheLysAsnValSerPhe 358
 DB 1160 AAGAGTGGACATAAACACAGATAATATTAAAGGAAATTTGGAATTTCAAAATGTTCACTTC 1219
 QY 359 AsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSer 378
 DB 1220 AGTTACCTCTCGAAAAGAAAGTTAAGATCTTAAAGGGTCTCAACCTGAAGGTTCAAGT 1279
 QY 379 GlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeu 398
 DB 1280 GGCACAGCTGCGCTGTTGGAAACAGTGGCTGCGGAAGACGACCGGTGACGCTG 1339
 QY 399 LeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArg 418
 DB 1340 ATGCAGAGCTCTATGACCCACAGATGCGCATGCTGTATTGATGGACAGGACATTAGG 1399
 QY 419 AlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeu 438
 DB 1400 ACCATAAATGTAAGGCATCTTCGGGAAATTTACTGGTGTGGTGGTGGTGGTGGTGGT 1459
 QY 439 PheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspValThrAspGluGlu 458
 DB 1460 TTTCACACACAGATAGCTGAAAACATTCCTATGCGCGGAAAATGTCACCATGGATGAG 1519
 QY 459 MetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLys 478
 DB 1520 ATTGAGAAAGCTTTAAGGAAGCCAAATGCTATGATTTTATCATGAACCTACCTAATAA 1579
 QY 479 PheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIle 498
 DB 1580 TTTCACACTCTGTTGGAGAGAGAGGGGCCAGCTGAGTGGTGGACAGAAACAGAAATC 1639
 QY 499 AlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSer 518
 DB 1640 GCATTTGCTGGGCCCTTGGTTTGCACCCCAAGATTTCTTCTGCTGATGAGGCAAGCTCA 1699
 QY 519 AlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
 DB 1700 GCTCTGCACACTGAAAGTGAAGCAGTGGTTTCAAGTGGCCCTGGATAAG 1747
 RESULT 8
 ID AAF86128 standard; cDNA; 4195 BP.
 XX AAF86128;
 AC AAF86128;
 XX 25-JUN-2001 (first entry)
 DE Cynomologous monkey P-glycoprotein cDNA variant 1.
 XX

Db 1363 AGTGGCAGACGGTCCCTGGTGGAAACAGCGGCTGTGGAGAGCAACACGGTCCAG 1422
 QY 398 LeuLeuGluArgLeuTyrAspProAspAspGlyPheLeuMetValAspGluAsnAspIle 417
 Db 1423 CTGATCGAGGCTTTATGACCCACAGAGGCATGTCAGTGTTCATGGACAGGATATT 1482
 QY 418 ArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProVal 437
 Db 1483 AGGACCAATAACCTAAGTTCTTACGGAAATCATCGTGTGTGAGTCAGGAACCTGTA 1542
 QY 438 LeuPheGlyThrThrLeuSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGlu 457
 Db 1543 TTGTTTCCACACCATAGTATGTAACCATTCGTATGTCGTGTCGTGTCATCATGGAT 1602
 QY 458 GluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsn 477
 Db 1603 GAGATTGAGAAAGCTGTCAAGGAAGCAATGCTATGACTTTATCATGAACCTGCCTCAG 1662
 QY 478 LysPheAsnThrLeuValGlyGluLysGlyValAlaGlnMetSerGlyGlyGlnLysGlnArg 497
 Db 1663 AATTTTCACACCTCGTTGGAGAGAGGGGCCCGCCAGCTGAGTGTGGCAGAGCAGAGG 1722
 QY 498 IleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThr 517
 Db 1723 ATCGCCATTGCAGTGCCTCGTTGTCGACCCCAAGATCCTCTGCTGGACGAGGCCAG 1782
 QY 518 SerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
 Db 1783 TCAGCCTTGGACACAGAAAGTGAAGCAGCTGTTTCAGTGGCTCTGGATAAG 1833
 RESULT 9
 AAZ49332
 ID AAZ49332 standard; cDNA; 3860 BP.
 XX AC AAZ49332;
 XX DT 14-MAR-2000 (first entry)
 XX DE Human wild-type multidrug resistance-1 (MDR-1) cDNA.
 XX KW Multidrug resistance; MDR-1; P-glycoprotein;
 KW transmembrane efflux pump; haematopoietic stem cell; transduction;
 KW bone marrow transplantation; chemotherapy; radiation therapy; cancer;
 KW gene therapy; gene replacement; genetic defect; thalassaemia;
 KW Gauchier's disease; sickle cell anaemia; leukaemia; ex vivo expansion;
 KW cytokine; wild-type; ds.
 XX OS Homo sapiens.
 XX FH Key
 XX CDS 1..3843
 FT Location/Qualifiers
 FT /tag= a
 FT /product= "Human wild-type MDR-1 protein"
 FT replace (553..555, GTR)
 FT /tag= b
 FT /note= "cDNA sequence of G185V human mutant MDR-1 given
 in AAZ49333"
 XX MO9961589-A2.
 XX PD 02-DEC-1999.
 XX 27-MAY-1999; 99WO-US11825.
 XX 28-MAY-1998; 98US-0086988.
 XX (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
 XX Sorrentino B, Bunting K;
 XX WFI; 2000-072615/06.
 XX P-PSDB; AAY58186.
 XX DR

PT Ex vivo expansion of hematopoietic stem cells transduced with a
 PT sequence encoding human multidrug resistance-1, used for bone marrow
 PT transplantation -
 PS Claim 10; Page 68-70; 113pp; English.
 XX This sequence represents cDNA encoding human wild-type
 CC multidrug resistance protein MDR-1. MDR-1 is a transmembrane
 CC efflux pump, responsible for the export of drugs from cells,
 CC particularly cancer cells. Wild-type MDR-1 shows increased
 CC resistance to etoposide and decreased resistance to vinca
 CC alkaloids compared with a mutant form (AAY58187) where the Gly at
 CC position 185 is replaced by Val. The invention relates to transducing
 CC haematopoietic stem cells with nucleic acid encoding an MDR protein
 CC and culturing the modified cells. The modified haematopoietic stem
 CC cells are useful in bone marrow transplantation (to reconstitute
 CC haematopoietic systems in patients who have undergone chemotherapy or
 CC radiation therapy) and in ex vivo gene therapy of genetic defects in
 CC cells derived from haematopoietic stem cells, e.g., thalassaemia,
 CC Gauchier's disease, sickle cell anaemia or leukaemia. The modified
 CC cells can also be used to identify factors involved in regulating
 CC proliferation and differentiation in haematopoietic stem cells.
 CC Haematopoietic stem cells that express MDR-1 will be protected against
 CC chemotherapeutic agents, so can be engrafted while the patient is
 CC undergoing chemotherapy. Expansion of (rare) haematopoietic stem cells
 CC provides sufficient cells to permit standard biochemical analysis.
 CC Overexpression of MDR-1 allows cytokine-driven expansion of
 CC haematopoietic stem cells by at least 10-fold compared with a maximum
 CC of 4-fold in known procedures.
 XX SQ Sequence 3860 BP; 1135 A; 746 C; 957 G; 1022 T; 0 other;
 Alignment Scores:
 Pred. No.: 2,52e-150 Length: 3860
 Score: 1555.50 Matches: 306
 Percent Similarity: 73.23% Conservative: 100
 Best Local Similarity: 55.23% Mismatches: 99
 Query Match: 56.79% Indels: 49
 DB: 21 Gaps: 6
 US-09-873-409-7 (1-541) x AAZ49332 (1-3860)
 QY 1 MetIleLeuGlyIleLeuAsnSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
 Db 151 ATGGTGGTGGAACTTTGGCTGCCATCATCCATGGGGTGGACTTCTCTCATGATGCTG 210
 QY 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
 Db 211 GTGTTTGGAGAAATGACAGATATCTTTGCCAATGCGAGGAAATTTAGAGATCTGATGTC 270
 QY 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
 Db 271 AACATCACTAATAGAGATGATATCATGATACAGGGTCTTTCATGATCTGGAGGAGAC 330
 QY 44 -----ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
 Db 331 ATGACACAGATATCGCTATTATTACAGTGAATTTGGTGTGGGGTGTGGTGTCTTAC 390
 QY 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleGlyGln 81
 Db 391 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAGAACAAATACACAAAATTAAGAAACAG 450
 QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
 Db 451 TTTTTCATGCTATATAGCAGAGATAGGCTGTTTGTATGTCACGATGTTGGGAG 510
 QY 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
 Db 511 CTTAACACCCGACTTACAGATGATGTCCTACAGATTAAATGAAGGTATTGGTGACAAAT 570
 QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
 Db 571 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGGATTTACACGT 630


```
QY 141 GlyTyrLysLeuThrValThrLeuSerThrSerProLeuLeuMetAlaSerAlaAla 160
DB 631 GGTGGAGCTAACCTTCTGATTTGGCCATCAGTCTCTGTTCTTGGACTGTGAGCTGT 690
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
DB 691 GTCTGGCAAGATACCTATCTCTATTTACTGTATAAAGAACTCTTAGCGTATGCAAAAGCT 750
QY 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
DB 751 GGAGCAGTAGCTGAAGAGCTCTGGCAGCAATTAGAATGTGATTCATTTGGAGGACAA 810
QY 201 GluLysGluLeuGlnArgSerPheLeuLeuAenIleThrArgTyrAlaTrpPheTyrPhe 220
DB 811 AAGAAGAACTTGAA----- 825
QY 221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrClnAenLeuLysAsp 240
DB 826 -----AGGTACCAACAAATAATTTAGAAAGAA 849
QY 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
DB 850 GCTAAAGAAATTGGGATAAAGAAAGCTATTACAGCCAAATATTTCTATAGTGTGCTTTC 909
QY 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAen 280
DB 910 CTGCTGATCTATGCATCTTATGCTCTGGCTTCTGCTATGGGACCACCTTGGTCTCTCA 969
QY 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
DB 970 GGGGAA-----TATTCTATTGGCAAGTACTCACTGTATTCTTTCTGTATTATTGGG 1023
QY 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
DB 1024 GCTTTAGTTGGACAGCACTCCAGCATTGAAGCAATTTGCCAATTCGAAGAGGACA 1083
QY 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
DB 1084 GCTTATGAATCTTCAAGATAATTGATAATAGCCAAATATTGACAGCTATTTCGAAGAGT 1143
QY 341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
DB 1144 GGGCACAACCAAGATAATATTAAAGGAAATTTGGAAATTCAGAAATGTTCACTTCAGTTAC 1203
QY 361 ProSerArgProSerIleLysIleLysGlyLeuAenLeuArgIleLysSerGlyGlu 380
DB 1204 CCATCTCGAAAGAAAGTTAAGATCTTGAAGGGCTGAACTGAAGGTGCAGAGTGGGCAG 1263
QY 381 ThrValAlaLeuValGlyLeuAenGlySerGlyLysSerThrValValGlnLeuLeuGln 400
DB 1264 ACGGTGGCCCTGGTTGGAAACAGTGGCTGGGAAGAGCACACAGCTCCAGCTGATGCGAG 1323
QY 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
DB 1324 AGGCTCTATGACCCACAGAGGGGATGGTTCAGTGTGGACAGGATATTAGGACCATA 1383
QY 421 AsnValArgHisTyrArgAspHisIleGlyValValSerClnGluProValLeuPheGly 440
DB 1384 AATGTAAGTTTCTACGGGAAATCATTTGGTGTGGTGCAGTCAGAAACCTGATTGTTGGC 1443
QY 441 ThrThrIleSerAsnAenIleLysTyrGlyArgAspAspValThrAspGluMetGlu 460
DB 1444 ACCACATAGCTGAAACATCTCGCTATGGCCGTGAAATGTACCATGGATGAGATTGAG 1503
QY 461 ArgAlaAlaArgGluAlaAenAlaTyrAspPheIleMetGluPheProAenLysPheAen 480
DB 1504 AAGCTGTCAAGGAAGCAATGCTATGCTATGCTGATGCTATGCTATGCTATGCTATGCT 1563
QY 481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyClnLysGlnArgIleAlaIle 500
DB 1564 ACCCTGGTTGGAGAGAGGGGGCCCGAGTTGAGTGGTGGGCGAAGAGAGGATTCGCCATT 1623
```

```
QY 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
DB 1624 GCACGTGCCCTGGTTGCGAACCCCAAGATCTCTCTGCTGATGAGCCACGTCAGCCTTG 1683
QY 521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
DB 1684 GACACAGAAGCGAAGCAGTGGTTGAGTGGTCTGGATAAG 1725
RESULT 10
ABA94365
ID ABA94365 standard; DNA; 3860 BP.
XX ABA94365;
XX
XX 26-MAR-2002 (first entry)
XX Human BCRP DNA related seq Id No. 1.
XX Stem cell; ATP transport protein; ATP-binding cassette; antiparkinsonian;
XX hepatotropic; neurodegenerative; cytotstatic; antianemic; muscular; BCRP;
XX cardiant; gene therapy; ds.
XX Homo sapiens.
FH Key Location/Qualifiers
FT CDS 1..3843
FT /*tag= a
XX
XX WO200192877-A2.
XX
XX PD 06-DEC-2001.
XX
XX PF 30-MAY-2001; 2001WO-US17459.
XX
XX PR 31-MAY-2000; 2000US-0584586.
XX
XX PR 29-MAY-2001; 2001US-0866866.
XX
XX (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX
XX PI Sorrentino B, Schuetz J;
XX
XX WPI: 2002-114368/15.
XX P-PSDB; ABB07266.
XX
XX Identifying a stem cell, for treating e.g., muscular dystrophy,
XX myocardial infarction, Parkinson's disease, or neurodegenerative
XX disorders, comprises detecting the expression of an ATP transport
XX protein (BCRP) by a cell
XX
XX Disclosure; Page 53-55; 87pp; English.
XX
XX The invention provides a method of identifying and/or isolating a stem
XX cell that involves detecting the expression of an ATP transport protein
XX containing a conserved ATP-binding cassette (BCRP) by a cell in a sample
XX comprising stem cells. The isolated stem cells may be used in the
XX treatment of diseases such as muscular dystrophy, degenerative liver
XX disorder, myocardial infarction, Parkinson's disease, degenerative
XX disorders of the brain, and for tissue regeneration or replacement.
XX Haematopoietic cells can be used in bone marrow transplants (e.g., for
XX treatment of leukemia) and for ex vivo gene therapy for treating blood
XX diseases such as sickle cell anemia and thalassemia. The stem cells can
XX also be used as cell targets in gene therapy protocols. The present
XX sequence represents a sequence related to the BCRP for which no relevant
XX information has been provided in the specification.
XX
XX SQ Sequence 3860 BP; 1135 A; 746 C; 957 G; 1022 T; 0 other;
```

Alignment Scores:

Pred. No.:	2,528-150	Length:	3860
Score:	1555.50	Matches:	306
Percent Similarity:	73.29%	Conservative:	100
Best Local Similarity:	55.23%	Mismatches:	99
Query Match:	56.79%	Indels:	49


```

DB:          24          Gaps:          6
US-09-873-409-7 (1-541) x ABA94365 (1-3860)

QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
DB 151 ATGGTGGTGGGAACCTTTGGCTGCGCATCATCCATGGGGCTGACCTCTCTCATGATGCTG 210
QY 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
DB 211 GTGTTTGGAGAAATGACAGATATCTTTGCAAAATGCGAGGAATTTAGAAATCTGATGTC 270
QY 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
DB 271 AACATCACTAATAGAAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 330
QY 44 -----ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
DB 331 ATGACCAGATATGCTTATTATTACAGTGGAAATTTGGTGGGGTGGCTGGTGTGCTGTAC 390
QY 62 IleGlnIleSerLeuTyrIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
DB 391 ATTACAGTTTCATTTGGTGGCTCGCGACGTGGAGCAAAATACACAAATTAGAAACAG 450
QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101
DB 451 TTTTTCATGCTATATATGCGACAGGAGATAGGCTGGTTTTCATGTCACGATGTTGGGAG 510
QY 102 LeuAsnThrArgMetThr-----AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
DB 511 CTTAAACACCCGACTTACAGATGATGCTCTTAAGATTAAATGAAGGTATTTGGTGACAAAT 570
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
DB 571 GGAATGTTCTTCAGTCAATGGCAACATTTTTCATCGGGTTTATAGTAGGATTTACACGT 630
QY 141 GlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
DB 631 GGTGGGAAGCTAACCTTGTGATTTTGGCCATCAGTCTGTTCTTGACGTGTCAGCTGCT 690
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
DB 691 GTCTGGGCAAGATACTATCTTCTTACTGTATAAGAACTCTTAGCGTATGCCAAAGCT 750
QY 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
DB 751 GGAGCAGTACTGAAGAGGCTTGGCGCAATTAGAATCTGTGATTCATTTGGAGACAA 810
QY 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTyrPheTyrPhe 220
DB 811 AAGAAAGAACTTGAA----- 825
QY 221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
DB 826 -----AGGTACACACAAATAATTTAGAGAA 849
QY 241 AlaIleAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
DB 850 GCTAAAGAAATTTGGGATAAGAGATTTACAGCCCAATATTTCTATAGGTGCTGCTTTC 909
QY 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrPyrGlyThrSerLeuIleLeuAsn 280
DB 910 CTGCTGATCTATGATCTTATGCTGCGCTTCTGGTATGGGACCACCTTTGGCTCTCA 969
QY 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
DB 970 GGGGAA-----TATTCTATTGGACAAAGTACTCACTGATTTCTTTCTGTATTAAATGGG 1023
QY 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
DB 1024 GCTTTTAGTGTGGACAGGATCTCCCAAGCATTTGAAGCATTTGCAATGCAAGAGGACGA 1083
QY 321 AlaPheHisIlePheGlnValIleAspLysPheProSerIleAspAsnPheSerThrAla 340

```

```

DB 1084 GCTTATGAAATCTTCAAGATAATTTGATAATAAGCCCAAGTATTTCACAGCTATTTCGAAGAGT 1143
QY 341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
DB 1144 GGGCACAAACCAAGATATATTAAAGGAAATTTTGGAAATTTGAGAAATGTTCACTTCAGTTAC 1203
QY 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
DB 1204 CCATCTCGAAAGAAAGTTAAGATCTTGAAGGCGCTGAACCTGAAGTGCAGAGTGGGCAG 1263
QY 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
DB 1264 ACGGTGCGCTGCTTGGAAACAGTGGCTGTGGAGAGACACACAGTCCAGCTGATGCAG 1323
QY 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
DB 1324 AGGCTCTATGACCCACAGAGGGGATGTCAGTGTGTGAGTGGACAGGATATTAGGACCAT 1383
QY 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
DB 1384 AATGTAAGGTTTCTACGGGAAATCATTTGCTGTGAGTCAGGAACTGTATTGTTTGC 1443
QY 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
DB 1444 ACCACGATAGCTGAAACATTCGCTATGCGCGTGAATAATGTCCACCATGGATGAGATTGAG 1503
QY 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
DB 1504 AAGCTGTCAAGGAAGCCCAATGCTATGATCTTATCATGAAACTGCTCATATAATTGAC 1563
QY 481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
DB 1564 ACCCTGTTGGAGAGAGAGGGCCCGCAGTTGAGTGTGGGAGAGCAGAGGATCGCCATT 1623
QY 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
DB 1624 GCACGTGCGCTGTTGCGAACCCCAAGATCTCTGCTGGATGAGGCCACGTGAGCCTTG 1683
QY 521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
DB 1684 GACACAGAAAGCGAAGCATGTTTTCAGGTGGCTCTGGATAAG 1725

RESULT 11
AAZ90198
ID AAZ90198 standard; DNA; 4233 BP.
XX
AC AAZ90198;
XX
DT 19-MAY-2000 (first entry)
XX
DE Rat mdrlb2 (multispecific drug transporter) nucleotide sequence.
XX
KW Rat; mdrlb2; multispecific drug transporter; cancer; inflammation;
KW cardiovascular disease; central nervous system disorder;
KW autoimmune disease; kidney disease; drug formulation; ss.
XX
OS Rattus sp.
XX
PN US6025160-A.
XX
PD 15-FEB-2000.
XX
PP 22-JUL-1998; 98US-0120513.
XX
PR 22-JUL-1998; 98US-0120513.
XX
PA (SMIK ) SMITHKLINE BEECHAM CORP.
XX
PI Brun KA, Ellens H, Yue L, Chenery RJ, Feild JA;
XX
DR WPI; 2000-181810/16.
DR P-PSDB; AAY78879.

```


XX Isolated polynucleotide encoding a rat mdrlb2 drug transporter
PT polypeptide, useful for treatment of e.g. cancer, autoimmune disease,
PT central nervous system disorders -
XX
PS Claim 7; Column 13-17; 17pp; English.

XX This sequence represents a polynucleotide encoding a 1275 amino acid
CC rat mdrlb2 multispecific drug transporter polypeptide. The mdrlb2
CC nucleotide sequence has cytostatic, antinflammatory, cardiant,
CC neuroprotective, immunosuppressive and nephrotropic activity.
CC understanding the functioning of the mdrlb2 polynucleotide and protein in
CC transgenic animal models is useful for treating and preventing diseases
CC such as cancer, inflammation, cardiovascular disease, central nervous
CC system disorders, autoimmune disease, and kidney disease. The use of the
CC protein in cell based, membrane based, or binding assays may enhance drug
CC formulation, selection of formulation excipients and compound design.

XX Sequence 4233 BP; 1210 A; 920 C; 1048 G; 1055 T; 0 other;

Alignment Scores:

Pred. No.: 5,88e-150 Length: 4233
Score: 1552.50 Matches: 308
Percent Similarity: 72.20% Conservative: 92
Best Local Similarity: 55.60% Mismatches: 105
Query Match: 56.68% Indels: 49
DB: 21 Gaps: 5

US-09-873-409-7 (1-541) x AAF290198 (1-4233)

Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 170 ATGGCTCTGGGAACCTCTGCTGCTATCATCCAGGAACTCTGCTCCCTCTCTGATGCTG 229
Qy 21 ValLeuGlyGluMetSerAspAsnLeuIle----- 30
Db 230 GTGTTCGGATACATGACATAGATGTTTACCACAGCAGACCCGCACTTCGCGGAGCGTT 289
Qy 31 -----SerGlyCysLeuValGlnThrAsnThrTySerPhePheArgLeu----- 45
Db 290 ACTAATCAAGTGAATCAACAGTACACAGCGCTCAGCAGCAGCTCGGAGGAGGAC 349
Qy 46 -----ThrLeuTyTyTyValGlyIleGlyValAlaAlaLeuIlePheGlyTy 61
Db 350 ATGGCCATGTACGCTACTATTACAGCGCATTTGGTGGCGTGTCTCATCTGTTGCTTAC 409
Qy 62 IleGlnIleSerLeuTrpIleThrAlaAlaArgGlnThrIleArgIleArgIleGln 81
Db 410 ATCCAGGTTTCACTTTGGTGCCTGGCGCTGGGAGACAAATACACAGATTAGGCGAGAG 469
Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyPheAspSerCysAspIleGlyGlu 101
Db 470 TTTTTCATGCCATCAATGATAGGAGATAGGCTGTTTGACGTGATGATGCGTGGGAG 529
Qy 102 LeuAsnThrArgMetThr---AspIleAspIleSerAspGlyIleGlyAspIle 120
Db 530 CTCACACCCGGCTCACAGATGACGCTCCAAAATAATGACGGAATTTGTCGACAACTT 589
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db 590 GGAATGTTCTTTCAGTCCCAATACGACATTTTTCAGCCGGTTTATAATAGGATTTAAGT 649
Qy 141 GlyTrpIleLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db 650 GGTGTGAAGCTAACCTTGTAATTTTGGCGCTCAGCCCTCTTATTGGTGTCTATCTGCC 709
Qy 161 AlaCysSerArgMetValIleSerLeuSerIleGlyLeuSerLeuSerIleSerLysAla 180
Db 710 ATGTGGGCAAGGACTGACTTCTTACTTAAATAGGAACCTCCAGGCTTATGCAAGCT 769
Qy 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db 770 GGAGCAGTTGCCGAGAGAGTCTTAGCAGCCATCAGAACTGTGATTGCGTTGGAGGACAA 829

Qy 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyAlaTrpPheTyPhe 220
Db 830 AAGAAGGAACCTTGA----- 844
Qy 221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyThrGlnAsnLeuLysAsp 240
Db 845 -----AGTACATATAAAATTTAGAGAA 868
Qy 241 AlaLysAspPheGlyIleLeuArgThrIleAlaSerLysValSerLeuGlyAlaValTy 260
Db 869 GCTAAAGAGTTGGCATTAAGAAACCCATCAGGCCAACATTTCCATAGGTATTCCTTAC 928
Qy 261 PhePheMetAsnGlyTyTyGlyLeuAlaPheTrpTyGlyThrSerLeuIleLeuAsn 280
Db 929 CTGTGGTCTATGCTCTATGCACTGGCATTCGTGTATGGACCTCCTTGTCTCTCTCA 988
Qy 281 GlyGlyProGlyTyThrIleGlyThrValLeuAlaValPheSerValIleHisSer 300
Db 989 AATGAA-----TATTCATTGGCAAGTGTCTTACCGTCTTCTCTATTTATTGGGG 1042
Qy 301 SerTyrcysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
Db 1043 ACTTTCAGTATTGGACATTTAGCCCCCAACATAGAGCCCTTTTGCAATTCAGAGGGCA 1102
Qy 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
Db 1103 GCCTATGAAATCTTCAAGATAATTGATAATGAGCAACCATTCAGACGCTTCTCAACCAAG 1162
Qy 341 GlyTyrcysProGluSerIleGlyThrValGluPheLysAsnValSerPheLeuTy 360
Db 1163 GGACACAAACAGACAGTATATATGGAAATTTGGAAATTTAAATAATGTTTACTTCAACTAC 1222
Qy 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
Db 1223 CCATCAGAGAGTGAAGTTAAGATCTTGAAGGCGCTCACTGAGGTGAAGGCGGCGAG 1282
Qy 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
Db 1283 ACGGTAGCCCTGTTGGCAACAGTGGCTGTGGGAAAGCACAACTGTCCAGCTGCTGCAG 1342
Qy 401 ArgLeuTyrcysProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
Db 1343 AGGCTCTACAGCCCATAGAGGGCGAGTTCAGTATCGACGGACAGGACATCAGACCATC 1402
Qy 421 AsnValArgHisTyrcysArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
Db 1403 AATGTGAGGTATCTCGGGGAATCATTTGGGTGGTGGTGTGAGTCAGGAACCCGCTGTTGCC 1462
Qy 441 ThrThrIleSerAsnAsnIleLysTyrcysArgAspValThrAspGluLeuMetGlu 460
Db 1463 ACCAGATTGCGGAAACATTCGCTATGCGCGAGAAACGTCACCATGATGATGATAGAG 1522
Qy 461 ArgAlaAlaArgGluAlaAsnAlaTyrcysPheIleMetGluPheProAsnLysPheAsn 480
Db 1523 AAGCTGTCAAGGAAGACCAATGCTATGATCTTCATCATGAAACTGCCCCACAAATTAAC 1582
Qy 481 ThrLeuValGlyLysGlyValGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
Db 1583 ACCCTGTTGGTGGAGAGGGCGGCGAGTGTGGTGGGAGACAGAAACAGAGGATCGCCATT 1642
Qy 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
Db 1643 GCCCGGCCCTGTTCGCAACCCCAAGATCTTTTGTGATGAGGCCACGTCAGCCTTG 1702
Qy 521 AspSerGluSerLysSerAlaValGlnAlaLeuGluLys 534
Db 1703 GACACAGAAAGGAGCGCTGTTTTCAGGCGCTCTCGGATAAG 1744

RESULT 12

AAF27498

ID AAF27498 standard; cDNA; 4233 BP.

XX


```

Db 1223 CCATCAGAGTGAAGTCTTGAAGGCGCTCAACTGAAGTGAAGAGCGGCGAG 1282
Qy 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuGln 400
Db 1283 ACGGTAGCCCTGGTGGCAACAGTGGCTGGGAAAGCACAACTGTCCAGCTGTGCGAG 1342
Qy 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
Db 1343 AGGCTCTACGACCCCATAGACGGCGAGGTTCAGTATCGACGACAGACATCAGGACCATC 1402
Qy 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGlnProValLeuPheGly 440
Db 1403 AATGTAGGTATCTCGGGAATCATTTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1462
Qy 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluMetGlu 460
Db 1463 ACCAGATTGCCGAAACATTCGCTATGCGCGAGAAAGCTCACCATGATGAGATAGAG 1522
Qy 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
Db 1523 AAAGCTGTCAAGGAAGCCAAATGCTATGACTTCATCATGAAACTGCCCCACAAATTAAAC 1582
Qy 481 ThrValGlyGlyLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
Db 1583 ACCCTGGTGGTGGAGAGGGGCGCGAGCTGAGTGGGGGACAGAAACAGAGGATCGCCATT 1642
Qy 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
Db 1643 GCCCGGCCCTGGTGGCAACCCCAAGATCCCTTTGTTGGATGAGGCCACGTCAGCCTTG 1702
Qy 521 AspSerGluSerLysSerAlaValGlnAlaLeuGluLys 534
Db 1703 GACACAGAAAGCGAGCGCTGGTTTCAGGCGCTCTCGATAAG 1744

RESULT 13
AAD03488
ID AAD03488 standard; cDNA; 4279 BP.
XX
AC AAD03488;
XX
DT 13-JUN-2001 (first entry)
XX
DE Dog P-glycoprotein (PGP) cDNA #1.
XX
KW Dog; P-glycoprotein; PGP; multidrug transporter; MDR1;
XX drug bioavailability; transgenic animal; genetic model; ss.
XX
OS Canis familiaris.
XX
FH Key Location/Qualifiers
FT CDS 17..3862
FT /*tag= a
FT /product= "Dog P-glycoprotein (PGP) #1"
XX
XX WO200123540-A2.
XX
PD 05-APR-2001.
XX
XX 28-SEP-2000; 2000WO-US26767.
XX
XX 28-SEP-1999; 99US-0156510.
XX
PA (GENT-) GENTEST CORP.
XX
PI Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;
XX
XX WPI; 2001-235373/24.
XX
DR P-PSDB; AAE00303.
XX
XX New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
XX for determining the bioavailability of drugs and for screening for dog
XX PGP inhibitors -

```

```

XX Claim 3; Page 58-63; 11lpp; English.
XX
CC The invention relates to dog P-glycoprotein (PGP) also referred
CC as multidrug transporter (MDR1) and nucleic acids encoding them.
CC The invention also includes fragments and biologically functional
CC variants of dog P-glycoprotein. PGP and their nucleic acids are
CC useful for determining the bioavailability of drugs and for
CC screening PGP inhibitors. They are useful for the diagnosis and
CC treatment of conditions characterised by PGP activity, by
CC reducing or increasing PGP activity in a cell. PGP nucleic acids
CC are used as oligonucleotide probes. Complements of PGP nucleic
CC acids are useful as antisense oligonucleotides, to induce a PGP
CC 'knockout' phenotype. They are used to prepare a non-human
CC transgenic animal, which are valuable as genetic models for
CC human diseases.
CC The present sequence is dog P-glycoprotein (PGP) cDNA. This
CC sequence is also referred as Genotype C cDNA. The
CC PGP enzyme functions as an efflux pump exporting small molecules
CC across the cell membrane. This enzyme is a member of the ABC
CC transporter family.
XX
SQ Sequence 4279 BP; 1294 A; 834 C; 1008 G; 1143 T; 0 other;

Alignment Scores:
Pred. No.: 5.97e-150 Length: 4279
Score: 1552.50 Matches: 302
Percent Similarity: 71.94% Conservative: 98
Best Local Similarity: 54.32% Mismatches: 105
Query Match: 56.68% Indels: 51
Db: 22 Gaps: 5

US-09-873-409-7 (1-541) x AAD03488 (1-4279)
Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyValCysLeuProLeuMetProLeu 20
Db 167 ATGTTGGTGGGACAAATGGCTGCCATCATCATTGAGAGTGCACCTCCCTCTCATGATGCTG 226
Qy 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu----- 34
Db 227 GTTTTGGAAACATGACAGATAGCTTTGCAATGCGAGGAATTTCAAGAAACAAACTTTT 286
Qy 35 -----ValGlnThrAsnThrTyrSerPhePheArg----- 44
Db 287 CCAGTTATAATTAATGAAAGTATTACGAACAATACACAACTTTTCATCAACCATCTGGAG 346
Qy 45 -----LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuPhe 59
Db 347 GAGGAAATGACCACGATATGCTTATTATACATGCGGATCGGTGCTGGCTGGTGGCT 406
Qy 60 GlyTyrIleGlnIleSerLeuTyrIleIleThrAlaAlaArgGlnThrLysArgIleArg 79
Db 407 GCTTACATCCAGTTTCATTCTGTCCTGGCAGCAGGAGACAGATACTCAAAATTAGA 466
Qy 80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIle 99
Db 467 AAACAATTTTTCATGCTATCATCGACAGAGATGGCTGGTTTCACCTGCATGACGTT 526
Qy 100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
Db 527 GGGGAGCTTAAACACCCGGCTCACAGCATGTCTCAAAATCAATCAAGGAATTCGCGAC 586
Qy 119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
Db 587 AAATTTGGAATGTTCTTTCACTCAATAGCAACATTTTTCACCGTTTATAGTGGGTTT 646
Qy 139 ValLysGlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
Db 647 ACACGTGGTGGGAAGCTAACCTTGTGATTTTGGCCATCAGCCCTGTTCTTGACATTCA 706
Qy 159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
Db 707 GCCGCCATCTGGGCAAGATACTACTTCTTCACTTACTATAAAGAACTCTTGGCCTATGCA 766

```


Qy	359	AsnTyrProSerArgProSerIleLeuLysGlyLeuAsnLeuArgIleLysSer	378
Db	1270	AGTTACCTCTCGAAAGAAGTAAGATCTTTAAAGGGTCTCAACTCAAGGTTACAGT	1329
Qy	379	GlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValcInLeu	398
Db	1330	GGGCAAGACAGTGGCGTGGTGGAAACAGTGGCTCGGGAGAGACGACCGCTGCAGCTG	1389
Qy	399	LeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArg	418
Db	1390	ATGCAGAGGGCTCTATCACCCACAGATGGCATGGTCTGTATTGATGGACAGACATTAGG	1449
Qy	419	AlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeu	438
Db	1450	ACCAATAATTAAGGCATCTTCGGGAAATTACTGTGTGGTGAGTCAGGAGCCTGTGTTG	1509
Qy	439	PheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGlu	458
Db	1510	TTTGGCCACCATGATAGCTGAAACATTCGCTATGATGCCGGAAATGTCCACCATGATGAG	1569
Qy	459	MetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLys	478
Db	1570	ATTGAGAAAGCTGTTAAGGAAGCCAATGCTATGATTTTATCATGAAACTACCTAATAAA	1629
Qy	479	PheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIle	498
Db	1630	TTTGACACTCTGGTTCGAGAGAGAGGGCCCGCTGAGTGGTGGGACAGAAACAGAGAATC	1689
Qy	499	AlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSer	518
Db	1690	GCCATTGCTGGGCCCTGGTTGCAACCCACAGATCTCTCTGCTGGATGAGCAACGTCA	1749
Qy	519	AlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys	534
Db	1750	GCTCTGGACACTGAAAGTGAAGCAGTGGTTCAGGTGGCCCTGGATAAG	1797
RESULT 15			
AAV32645			
ID	AAV32645 standard; cDNA; 4669 BP.		
XX	AAV32645;		
XX	23-SEP-1998 (first entry)		
XX	Human P glycoprotein (Pgp) cDNA.		
XX	Human P glycoprotein; Pgp; multi-drug resistance; cancer;		
KW	UIC2 monoclonal antibody; mAB; cytotoxic; transmembrane efflux pump; s		
XX	Homo sapiens.		
XX	Key		
FH	Location/Qualifiers		
FT	5'UTR		
FT	1..424		
FT	/*tag= a		
FT	CDS		
FT	425..4267		
FT	/*tag= b		
FT	/product= "wild-type Pgp"		
FT	4268..4669		
FT	/*tag= c		
XX	WO9821325-A1.		
XX	22-MAY-1998.		
XX	17-NOV-1997; 97WO-US21214.		
XX	15-NOV-1996; 96US-0752447.		
XX	(INGE-) INGENEX INC.		
PA	(UNII) UNIV ILLINOIS FOUND.		
XX	Mechetner E, Roninson IB;		

XX WPI: 1998-297930/26.
 DR P-PSDB; AAW48997.
 XX
 PT Immunological reagent specific for P-glyco:protein - useful for
 PT detecting multi-drug resistant cancer, isolating haematopoietic
 PT cells and selective cell killing
 XX
 PS Claim 1; Fig 1A; 89pp; English.
 XX
 CC The present sequence represents the human P glycoprotein (Pgp)
 CC CDNA which encodes the wild-type Pgp protein. Pgp is a transmembrane
 CC efflux pump protein involved in multi-drug resistance of cancer cells.
 CC The invention provides methods for developing and using immunological
 CC reagents specific for certain mutant forms of Pgp and wild-type Pgp in
 CC a conformation associated with substrate binding or in the presence
 CC of ATP depleting agents. An example of the immunological reagent
 CC is the UIC2 monoclonal antibody (mAb). mAb UIC2 specifically binds
 CC to Pgp in a particular biochemical conformation and is capable of
 CC inhibiting drug efflux from Pgp-expressing cells. The immunological
 CC reagents are claimed to be useful for detecting Pgp expression in
 CC mammalian cells, including low level expression, particularly in cancer
 CC cells to diagnose multi-drug resistance. The invention claims that
 CC these immunological reagents are more specific than known reagents for
 CC detecting Pgp and they also eliminate the need for costly and laborious
 CC screening of Pgp inhibitors by cytotoxicity or dye exclusion methods.
 XX
 SQ Sequence 4669 BP; 1393 A; 894 C; 1130 G; 1252 T; 0 other;

Alignment Scores:
 Pred. No.: 6.77e-150 Length: 4669
 Score: 1552.50 Matches: 305
 Percent Similarity: 73.29% Conservative: 101
 Best Local Similarity: 55.05% Mismatches: 99
 Query Match: 56.68% Indels: 49
 DB: 19 Gaps: 6

US-09-873-409-7 (1-541) x AAV32645 (1-4669)

QY 1 MetileuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
 DB 575 ATGGTGGGAACTTTGGTGGCCATCATCCATGGGGCTGGAGTTCCTCTCATGATGCTG 634
 QY 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
 DB 635 GTGTTGGAGAAATGACAGATATCTTGCRAATGCAGGAAATTAGAGATCTGATGCA 694
 QY 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
 DB 695 AACATCAATAAGATGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 754
 QY 44 -----ArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
 DB 755 ATGACCAAGGTATGCTATTATATACAGTGAATGGTGGGGTGGTGGTGGTGGTGGTTC 814
 QY 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
 DB 815 ATTCAAGTTTCATTGTTGGTGGCTGGCAGCTGGAAGACAAATACACAAATTAGAAACAG 874
 QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
 DB 875 TTTTTCATGCTATTATCGGACAGGAGATAGCTGGTTGATGTGCACGATGTTGGGAG 934
 QY 102 LeuAsnThrArgMetThr---AspIleAspIleSerAspGlyIleGlyAspIle 120
 DB 935 CTTAACACCCGACATTACAGATGATGCTCCCAAGATTAATGAAGGAATGGGACAAAT 994
 QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
 DB 995 GGAATGTTCTTTCAGTCAATGGCACATTTTTCAGTGGTTTATAGTAGGATTTACAGT 1054
 QY 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160

DB 1055 GGTGGAAAGTAACCTTGTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTGAGCTGCT 1114
 QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
 DB 1115 GTCTGGGCAAGATACATACTTCTTCACTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 1174
 QY 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
 DB 1175 GGAGCAGTAGCTGAGAGAGTCTTGGCAGCAATTAGAAGCTGTGATTCATTTGGAGACAA 1234
 QY 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
 DB 1235 AGAAAGAACTTGA----- 1249
 QY 221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
 DB 1250 -----AGGTACAACAACAAATTTAGAGAA 1273
 QY 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
 DB 1274 GCTAAAAGAAATTTGGGATAAAGAAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTC 1333
 QY 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
 DB 1334 CTGCTGATCTATGCACTTATGCTGGCCCTTCTGGTATGGACACCTTGGCTCTCA 1393
 QY 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
 DB 1394 GGGGAA-----TATTCATTGGCAAGTACTCACTGTATTCTTTCTGTATTAAATTTGGG 1447
 QY 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAla9GlyAla 320
 DB 1448 GCTTTTAGTGTGGACAGGCATCTCCAAGCATTTGAAGCATTTGCAAAATGTCAGAGGAGCA 1507
 QY 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
 DB 1508 GCTTATGAATCTTCAAGATAATTGATAATGACCAAGTATTGACACTATTGCAAGAGT 1567
 QY 341 GlyTyrLysProGluSerIleGlyThrValGluPheLysAsnValSerPheAsnTyr 360
 DB 1568 GGGCACAAACCATGATAATTAAGGGAAATTTGGAATTCAGAAATGTTCACTTCACTTAC 1627
 QY 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
 DB 1628 CCATCTCGAAAAGAAAGTTAAGATCTTGAAGGGCTGAAACCTGAAGGTGCAGAGTGGGCAG 1687
 QY 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
 DB 1688 ACGTGGCCCTGGTTGGAAACAGTGGCTGTGGGAAGACACACAGTCCAGCTGATGTCAG 1747
 QY 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
 DB 1748 AGSCTCTATACCCACAGAGGGGATGGTGCAGTGTTCATGTCAGCAGGATATTAGACCATA 1807
 QY 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
 DB 1808 AATGTAAGTTCCTACGGGAAATCATTTGGTGTGTGTGAGTCAGGAACCTGATTGTTGGC 1867
 QY 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
 DB 1868 ACCACGATATCGTGAACCATTTGCTATGGCCGGTGAATGTCCACTGATGAGATTGAG 1927
 QY 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
 DB 1928 AAGCTGTCAAGNAGCAATGCTATGACTTTATCATGAACTGCCTCATAAATTTGAC 1987
 QY 481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
 DB 1988 ACCCTGTTGGAGAGAGAGGGGCCAGTGTGAGTGTGGGAGAGAGCAGAGGATGCCCAT 2047
 QY 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
 DB 2048 GCAGTGGCTGGTTGCAACCCCAAGATCTCTCTGCTGATGAGGCCACGCTCAGCCTTG 2107

Qy 521 AspSerGluSerIysSerAlaValGlnAlaLeuGluLys 534
|||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 2108 GACACAGAAAGCGAAGCACTGGTTTCAGGTGGCTCTGGATAAG 2149

Search completed: March 30, 2003, 03:21:35
Job time : 587.737 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4 p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 30, 2003, 03:22:44 ; Search time 3721.1 Seconds
(without alignments)
2354.618 Million cell updates/sec

Title: US-09-873-409-7
Perfect score: 2739
Sequence: 1 MILGILASLVNGACILPLMLP.....SEKSAVQAALKEKTPRYSP 541

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-O=/cgn2_1/USPTO.spool/US09873409/runat_27032003_115420_19240/app_query.fasta_1.7544
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09873409 @CGN 1 1 30544 @runat_27032003_115420_19240 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hc:*
9: gb_estcl:*
10: gb_est2:*
11: gb_hc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1092	39.9	669	13	BM013981	BM013981 603639619
2	1068.5	39.0	1033	14	BM926413	BM926413 AGENCOURT
3	1029	37.6	760	13	BM016204	BM016204 603642659
4	918	33.5	545	9	AL040762	AL040762 DKP29434C
5	903	33.0	929	14	BQ882401	BQ882401 AGENCOURT
6	856.5	31.3	948	13	BM471690	BM471690 AGENCOURT
7	750	27.4	894	12	BF584668	BF584668 602098406
8	744.5	27.2	872	11	AK020318	AK020318 Mus muscu
9	689.5	25.2	2676	11	AK014319	AK014319 Mus muscu
10	659	24.1	943	9	AL520322	AL520322 AL520322
11	654	23.9	405	9	AA243820	AA243820 zr67906.r
12	650	23.7	550	13	BJ075336	BJ075336
13	648	23.7	803	12	BG298756	BG298756 602396681
14	637	23.3	780	12	BG587938	BG587938 EST489713
15	625	22.8	559	12	BG692596	BG692596 602248949
16	621	22.7	606	14	BQ262340	BQ262340 fz85a08.x
17	621	22.7	681	14	BQ06064	BQ06064 QGG9N16.Y
18	604.5	22.1	932	17	AZ670821	AZ670821 ENTJUN69TF
19	602.5	22.0	913	17	BH155700	BH155700 ENTRO54TR
20	599	21.9	880	17	AZ687805	AZ687805 ENTLU52TF
21	596.5	21.8	871	17	AZ682350	AZ682350 ENTKB16TF
22	590.5	21.6	891	17	AZ682250	AZ682250 ENTKT68TR
23	585.5	21.4	853	17	AZ679807	AZ679807 ENTHI67TR
24	583	21.3	2564	11	AK020022	AK020022 Mus muscu
25	582.5	21.3	787	12	BG584063	BG584063 EST485823
26	581	21.2	586	13	BJ336461	BJ336461
27	579	21.1	661	12	BG302675	BG302675 fl36db04.x
28	578	21.1	575	10	AW342644	AW342644 fl81f06.x
29	577.5	21.1	886	17	AZ540627	AZ540627 ENTQ18TF
30	577	21.1	835	13	BJ349604	BJ349604 BJA49604
31	576.5	21.0	947	17	AZ683753	AZ683753 ENTIL96TF
32	576	21.0	547	13	BJ333166	BJ333166
33	574.5	21.0	885	17	AZ686798	AZ686798 ENFPZ28TR
34	574	21.0	542	13	BJ339421	BJ339421
35	572.5	20.9	860	17	AZ543293	AZ543293 ENTQ80TF
36	572.5	20.9	886	17	BH139685	BH139685 ENTNA47TR
37	571	20.8	715	14	BQ869512	BQ869512 QGD6G09.Y
38	570	20.8	814	13	BJ352204	BJ352204
39	570	20.8	820	13	BJ356540	BJ356540
40	570	20.8	1019	12	BG248052	BG248052 602359987
41	569	20.8	1341	11	AY108285	AY108285 Zea mays
42	565	20.6	657	14	BQ802058	BQ802058 WHE2821.F
43	564	20.6	636	13	BJ386538	BJ386538
44	564	20.6	750	13	BJ445883	BJ445883
45	562	20.5	1813	11	U66688	U66688 Homo sapien

ALIGNMENTS

RESULT 1
BM013981
LOCUS 603639619P1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:5415556 5',
DEFINITION mRNA sequence.
ACCESSION BM013981
VERSION BM013981.1 GI:16528335
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (Bases 1 to 869)
AUTHORS NIH-MGC <http://hgnc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL
COMMENT

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12060 row: f column: 05
High quality sequence stop: 664.
Location/Qualifiers

FEATURES

Source
1..669
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5415556"
/clone_lib="NIH MGC 87"
/tissue type="mammary adenocarcinoma, cell line"
/lab host="DH10B (phage-resistant)"
/note="Organ: breast; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.383 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH MGC library."
BASE COUNT 203 a 122 c 162 g 182 t
ORIGIN
Alignment Scores:
Pred. No.: 2,18e-123 Length: 669
Score: 1092.00 Matches: 215
Percent Similarity: 99.54% Conservative: 2
Best Local Similarity: 98.62% Mismatches: 0
Query Match: 39.87% Indels: 1
DB: 13 Gaps: 0
US-09-873-409-7 (1-541) x BM013981 (1-669)

QY 292 AlavalPheSerValIleHisSerSerTyrCysIleGlyAlaValProHisPhe 311
DB 17 AGTGTTTCTTAATGTAATCCATAGCAGTATTGCGTTCGAGCAGGAGCCCTCACTTT 76
QY 312 GluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLys 331
DB 77 GAACCTTCGCATAGCCGAGGAGCTGCTTTTCATATTTTCCAGGTATTATAGAAA 136
QY 332 ProSerIleAspAspSerThrAlaGlyTyrLysProGluSerIleGluGlyThrVal 351
DB 137 CCCAGTATAGATAACTTTTCCACAGCTGGATATAAACCTGAATCCATAGAGAACTGTG 196
QY 352 GluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGly 371
DB 197 GAATTTTAAAAATGTTCTTTCAATATTATCCATCAAGACCATCTATCAAGATTCGAAAGGT 256
QY 372 LeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGly 391
DB 257 CTGAATCTCAGATTAGTCTGGAGACACAGTCGCTTGGTTCATGCGAGTGGG 316
QY 392 LysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMet 411
DB 317 AAGAGTACGGTAGTCCAGCTTCTGCAGAGGTATATGATCCGATGATGCTTATCATG 376
QY 412 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 431
DB 377 GTGGATGAGAAATGACATCAGAGCTTTTAAATGTGGGCAATTATCGAGACCATATTGGAGTG 436
QY 432 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 451
DB 437 GTTAGTCAAGAGCCCTGTTTGTTCGGACACCACTCATGTAACATATCAAGTATGGACGA 496
QY 452 AspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 471
DB 497 GATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGAGGAAGCAAAATGCGTATGATTTT 556

QY 472 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGlyAlaGlnMetSer 491
DB 557 ATCATGAGAGTTTCTTAATAAATTTATATACATTTGGTAGGGGAAAAAGAGCTCAATGAGT 616
QY 492 GlyGlyGlnLysGlnArg-IleAlaIleAlaArgAlaLeuValArgAsnPro 508
DB 617 GGAGGGCAGAAACAGAGGCGATCGCAATTGCTCGTGCTTTAGTTTCGAACCC 668

RESULT 2

BM926413
LOCUS
DEFINITION
5', mRNA sequence.
AGENCOURT 6600788 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5764845
BM926413
VERSION
BM926413.1 GI:19376792
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens

BM926413
VERSION
BM926413.1 GI:19376792
KEYWORDS
EST.
SOURCE
human.

ORGANISM
Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1033)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov

Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

Plate: LLAM12818 row: o column: 22
High quality sequence stop: 3
High quality sequence stop: 686.
Location/Qualifiers
1..1033
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5764845"
/clone_lib="NIH MGC_114"
/lab host="DH10B"
/note="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27 yo. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 019. Note: this is a NIH_MGC Library."

BASE COUNT 289 a 212 c 218 g 312 t
ORIGIN

Alignment Scores:
Pred. No.: 3,46e-120 Length: 1033
Score: 1068.50 Matches: 229
Percent Similarity: 93.11% Conservative: 3
Best Local Similarity: 91.97% Mismatches: 17
Query Match: 39.01% Indels: 4
DB: 14 Gaps: 0
US-09-873-409-7 (1-541) x BM926413 (1-1033)

QY 293 ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaValProHisPheGlu 312
DB 180 GTTTTCTTAGTGTATCCATAGCAGTATTGATGGAGCAGAGTCCCTCACTTTGAA 239

QY 313 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPro 332
DB 240 ACCTTCGCAATAGCCCGAGGAGTGGCTTTTCATATTTTCCAGGTTCATTGATAGAAACCC 299


```

QY 333 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 352
DB 300 AGTATAGATAAATCTTCCACAGCTGGATATAAACCTGGAATCCATAGAGAACTGTGGAA 359
QY 353 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLysGlyLeu 372
DB 360 TTTAAATAAGTTCTTCAATATATCCATCAAGACCATCTATCAAGATCTCGAAGGTCGT 419
QY 373 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 392
DB 420 AATCTCAGAAATTAAGTCTCGAGAGACAGTCGCTTGGTGGTCTCAATGCGAGTGGGAAG 479
QY 393 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 412
DB 480 AGTACGGTAGTCAGCTTCTCGAGAGGTTATATGATCCGGATGATGGCTTTATCATGGTG 539
QY 413 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal 432
DB 540 GATGAGATGACATCAGACTTTAAATGTGCGGCATATTCGAGACCATATTCGAGTGGTT 599
QY 433 SerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAsp 452
DB 600 AGTCAAGAGCCTGTTTGTTCGGGACCACCATCAGTAACATATCAAGTATGACGAGAT 659
QY 453 AspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIle 472
DB 660 GATGTGACTGATCAAGAGATGGAGAGAGCAGCAAGCAAGGAAGCAAAATCGGTATGATTATC 719
QY 473 MetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGly 492
DB 720 ATGGAGTTTCTTAATAA- TTTAATACATTGGTAGGGGAANAAGGAGCTCCCAATGATGGA 778
QY 493 GlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIle 512
DB 779 AGCCAGAACCG- AGGATCCCATTTGCTCTGCTTGTAGTTTGGAAACCCAGGAT- CTGAAT 836
QY 513 LeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaLeu 532
DB 837 TTAGATGAAGCTACGCTGCTCCCTGGATTGAGAAACCAAGTCAACTGTGTCA- GCTGGCCTG 894
QY 533 GluLysAspThrProArgTyrSerPhe 541
DB 895 GAGAAGGAAACCCCGGNAATTCCTTT 921

RESULT 3
BM016204 603642659F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:5418615 5',
LOCUS mRNA sequence. linear EST 30-OCT-2001
DEFINITION
ACCESSION BM016204
VERSION BM016204.1 GI:16530558
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgi.nci.nih.gov/
JOURNAL 1 (bases 1 to 760)
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-roman@nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Place: LLNL2068 row: e column: 16
High quality sequence stop: 740.
Location/Qualifiers

```

FEATURES

DEFINITION DKE2p434C1815_r1 434 (synonym: htes3) Homo sapiens cDNA clone

LOCUS AL040762

545 bp mRNA linear

EST 29-FEB-2000

source

1..760
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5418615"
 /clone_lib="NIH_MGC_87"
 /tissue_type="mammary adenocarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: breast; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 1.383 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."
 BASE COUNT 229 a 150 c 187 g 193 t 1 others
 ORIGIN

Alignment Scores:

Pred. No.: 1,52e-115 Length: 760
 Score: 1029.00 Matches: 209
 Percent Similarity: 98.12% Conservatives: 0
 Best Local Similarity: 98.12% Mismatches: 2
 Query Match: 37.57% Indels: 2
 DB: 13 Gaps: 0

US-09-873-409-7 (1-541) x BM016204 (1-760)

QY 330 LysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGly 349
 DB 3 AAGAAACCCAGTATAGATAAATCTTCCACAGCTGGATATAAACCTGGAATCCATAGAGGA 62
 QY 350 ThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeu 369
 DB 63 ACTGTGGAAATTTAAATAATGTTCTTCAATATATCCATCAAGACCATCTATCAAGATTCG 122
 QY 370 LysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGly 389
 DB 123 AAAGTCTGAATCTCAGAATTAAGTCTGGAGAGACAGTCGCTTGGTCTCAATGTC 182
 QY 390 SerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPhe 409
 DB 183 AGTGGGAAGAGTAGCGGTAGTCAGCTTCTCGCAGAGGTTATATGATCCGATGATGCTTT 242
 QY 410 IleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIle 429
 DB 243 ATCATGGTGGATGAGATGACATCAGAGCTTTAAATGTGCGGCATTTATCGAGACCATATN 302
 QY 430 GlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyr 449
 DB 303 GGAGTGGTTAGTCAGAGGCTGTTTGTTCGGGACCACCATCAGTAACATATCAAGTAT 362
 QY 450 GlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyr 469
 DB 363 GGACGAGATGATGTGACTGATCAAGAGATGGAGAGCAGCAAGGAAGCAAAATGCGTAT 422
 QY 470 AspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGln 489
 DB 423 GATTTTATCATGAGCTTTCTTAATAATTTAATACATTTGGTCTGGGGAAGAGAGCTCAA 482
 QY 490 MetSerGlyGlyGlnLysGlnArgIleAlaIle-AlaArgAlaLeuValArgAsnProLys 509
 DB 483 ATGAGTGGAGGAGCAAAACAGAGGATCGCAATGGCTCGTCTAGTTTCGAAACCCCAA 542
 QY 509 sileLeuIleLeu-AspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValG 529
 DB 543 GATTTCTGATTTTCAGATGAGGCTAGCTCTGCCCTGGATTCAGAAAGCAAGTCAGCTGTT 602
 QY 529 InAlaAlaLeuGluLysAspThrProArgTyrSer 540
 DB 603 AAGCTGCATCTGGAGAAGGATACCCCGAGGTATTC 637

RESULT 4

AL040762

545 bp mRNA linear

EST 29-FEB-2000

DKFZp434C1815 5', mRNA sequence.

AL040762
VERSION
AL040762.1 GI:5409708
EST.
SOURCE
ORGANISM
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 545)
Bloeker, H., Boecker, M., Brandt, P., Mewes, H.W., Gassenhuber, J. and
Wiemann, S.
EST (Bloeker, et al.)
Unpublished (1999)
Contact: Bloeker H
MIPS

Am Klopfergspitz 18a D-82152 Martinsried, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by GBF (National Research Centre for Biotechnology Ltd.,
Braunschweig/Germany) within the cDNA sequencing consortium of the
German Genome Project.
SI sequence also available.
This clone (DKFZp434C1815) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
Location/Qualifiers
1. .545
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZp434C1815"
/clone_lib="434 (synonym: htes3)"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"

BASE COUNT 171 a 89 c 138 g 147 t
ORIGIN

Alignment Scores:
Pred. No.: 48-102 Length: 545
Score: 918.00 Matches: 181
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 33.52% Indels: 0
DB: 9 Gaps: 0

US-09-873-409-7 (1-541) x AL040762 (1-545)

QY 333 SerIleAspAsnPhSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 352
Db 3 AGTATAGATCACTTTCCACAGCTGGATATTAACCTGATTCATGAGGAAGAACTGTGGA 62

QY 353 PhelyAsnValSerPheAntyProSerArgProSerIleLysIleLeuLysGlyLeu 372
Db 63 TTTAAAATGTTCTTCAATTCATTCATCAAGACCATCTATCAAGATTCGAAAGTCTG 122

QY 373 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 392
Db 123 AATCTCAGAAATTAAGTCGAGAGACAGTCGCTTGGCTCTCAATGGCAGTGGGAAG 182

QY 393 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspGlyPheIleMetVal 412
Db 183 AGTACGGTAGTCACGCTTCGACAGAGGTATATGATCCGGATGATGCTTTATCATGTG 242

QY 413 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal 432
Db 243 GATGAGATGACATCAGAGCTTTAAATGTGCGCATTCATCGACCATATTCGAGTGGTT 302

QY 433 SerGlnGluProValLeuPheGlyThrThrIleSerAsnIleLysTyrGlyArgAsp 452
Db 303 AGTCAAGAGCGCTGTTTGTTCGGGACCACCACCATCAGTAACAATATCATGATGGACGAG 362

QY 453 AspValThrAspGluGluMetGluArgAlaAlaAaGluAlaAaAlaTyrAspPheIle 472
Db 363 GATGTGACTGATGAAGAGATGAGAGAGCAGCAGGAAGAAATGCGTATGATTTATC 422

QY 473 MetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGlyValAlaGlnMetSerGly 492
Db 423 ATGGAGTTTCCTAATAAATTTAATACATTTGGTGGGAAAAGAGAGCTCAATGAGTGA 482

QY 493 GlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 512
Db 483 GGSCAGAAAACAGAGGATCGCAATTCCTGCTGCTAGTTCGAAACCCCAAGATTCGATT 542

QY 513 Leu 513
Db 543 TTA 545

RESULT 5
BQ882401
LOCUS
DEFINITION
BQ882401
VERSION
BQ882401.1 GI:22274409
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 929)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2492 row: c column: 15
High quality sequence stop: 677.
Location/Qualifiers
1. .929
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6291782"
/clone_lib="NIH_MGC_43"
/tissue_type="normal pigmented retinal epithelium"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC library. |"
BASE COUNT 281 a 176 c 237 g 234 t
ORIGIN

Alignment Scores:
Pred. No.: 6.7e-100 Length: 929
Score: 903.00 Matches: 188
Percent Similarity: 73.02% Conservative: 61
Best Local Similarity: 55.13% Mismatches: 58
Query Match: 32.97% Indels: 34
DB: 14 Gaps: 3

US-09-873-409-7 (1-541) x BQ882401 (1-929)

QY 176 AlaTyrSerLysAlaGlyAlaValAlaGluValSerSerIleArgThrValI 195


```
Db 1 CGGTATGCAAAAGCTGGACAGTAGCTGAAGAGGCTTGGCAGCAATTAGAAGCTGTGAT 60
Qy 195 eAlaPheArgAlaGlnGlnGluGlnGlnArgSerPheLeuLeuAsnIleThrArgTy 215
Db 61 TGCATTGGAGCACAAGAGAACTTGA----- 91
Qy 215 rAlaTrpPheTyPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyTh 235
Db 92 -----AGGTACAA 99
Qy 235 rGlnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSe 255
Db 100 CAAAATTTAGAGAGCTAAAGAAATTCGGATAAAGAAAGCTATTACAGCCCAATATTTC 159
Qy 255 rLeuGlyAlaValTyPhePheMetAsnGlyThrTyGlyLeuAlaPheTyrTyGlyTh 275
Db 160 TATAGTGTCTTCTCTCTCTCTATCATCTATCATCTATCTCTGGCCCTCTGGTATGGGAC 219
Qy 275 rSerLeuIleLeuAsnGlyGluProGlyTyThrIleGlyThrValLeuAlaValPhePh 295
Db 220 CACCTTGGTCTCTCAGGGGAA-----TATTCTATTGGACAAAGTACTCACTGATTCTT 273
Qy 295 eSerValIleHisSerSerTyrcysIleGlyAlaAlaValProHisPheGluThrPheAl 315
Db 274 TTCTGTATTAATTGGGGCTTTTAGTGTGGACAGGCATCTCCAAGCATTGAAGCATTTGC 333
Qy 315 alIeAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAs 335
Db 334 AATGCAAGAGGAGCGCTTATGAAATCTTCAGATTAATTGATAATGAACCAAGTATTGA 393
Qy 335 pAsnPheSerThrAlaGlyTyTyLysProGluSerIleGlyThrValGluPheLysAs 355
Db 394 CAGCTATTGGAAGAGTGGGCACAAACAGATAATAATTAAGGGAAATTTGGAATTCAGAAA 453
Qy 355 nValSerPheAsnTyTrpSerArgProSerLysIleLysIleLeuLysGlyLeuAsnLeuAr 375
Db 454 TGTTCACCTTCAGTTACCCATCTCGAAAGAAGTTAAGATCTTTGAAGGGTCTGAACCTGAA 513
Qy 375 rIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrVa 395
Db 514 GGTGCAGATGGGCAGACGGTGGCCCTGGTTGGAAACAGTCAGTGGCTGTGGGAAGAGCAAC 573
Qy 395 lValGlnLeuLeuGlnArgLeuTyTrpAspProAspAspGlyPheIleMetValAspGluAs 415
Db 574 AGTCCAGCTGTATGCAGAGCTCTATGACCCACAGAGGGGATGGTCAGTGTGATGGACA 633
Qy 415 nAspIleArgAlaLeuAsnValArgHisTyArgAspHisIleGlyValValSerGlnGl 435
Db 634 GGATATTAGGACCATAATGTAAGGTTTCTACGGGAAATCATTTGGTGTGGTGAAGTCA 693
Qy 435 uProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyTyGlyArgAspValTh 455
Db 694 ACTGTATTGTTTGGCCACACAGATAGCTGAAACATTCCTATGGCCGTGAAATGTGAC 753
Qy 455 rAspGluGluMetGluArgAlaAlaArgGluAlaAsnAla-TyrAspPheIleMetGluP 475
Db 754 CATGGATGAGATTGAAAGAGCTCTCAGGAGGCCAATGCCCTATGATCTTATCATGAAC 813
Qy 475 heProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGly--G 494
Db 814 TGCCTCATAAATTGACACCTGGTTGGANAAAGAGGGGCCCAATTGAGTGTGGGGGCAC 873
Qy 494 lNlysGlnArgIleAlaIleAlaArg-AlaLeuVal-ArgAsnProLysIleLeu 511
Db 874 AAACCAAGAGATCTCCCTTGGCAGTCGTGGCCCTGGTTCCCAACCCCAAGATCCTC 928

RESULT 6
LOCUS BM471690
DEFINITION AGENCOURT_6465349 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5539117
5', mRNA sequence.
ACCESSION BM471690
```

```
VERSION BM471690.1 GI:18520732
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 948)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC/DCID/DTP
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Disseminating by: Agencourt Bioscience Corporation
Clone sequencing: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12233 row: b column: 14
High quality sequence stop: 569.
FEATURES
Location/Qualifiers
1..948
source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5539117"
/clone_lib="NIH_MGC_72"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: skin; Vector: PCMV-SPORT6; Site 1: Not I;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies."
BASE COUNT 285 a 187 c 194 g 282 t
ORIGIN
Alignment Scores:
Pred. No.: 3,646-94 Length: 948
Score: 856.50 Matches: 186
Percent Similarity: 89.67% Conservative: 5
Best Local Similarity: 87.32% Mismatches: 14
Query Match: 31.27% Indels: 8
DB: 13 Gaps: 1
US-09-873-409-7 (1-541) x BM471690 (1-948)
Qy 293 ValPhePheSerValIleHisSerSerTyrcysIleGlyAlaAlaValProHisPheGlu 312
Db 205 GTTTCTTTTAGTGAATCATACAGCAGTTATTGCATTGGAGCAGCAGTCCCTCACTTTGAA 264
Qy 313 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPro 332
Db 265 ACCTTGCATAGTACCCGAGAGCTGCTTTTCATATTTTCCAGGTATTGATAAGAAACCC 324
Qy 333 SerIleAspAsnPheSerThrAlaGlyTyTyLysProGluSerIleGlyThrValGlu 352
Db 325 AGTATAGATAACTTTTCCACAGCTGGATATAAACCTTGAATCCATAGAGAACTGTGGAA 384
Qy 353 PheLysAsnValSerPheAsnTyTrpSerArgProSerIleLysIleLeuLysGlyLeu 372
Db 385 TTTAAATAATGTTCTTTCATATTCATCAAGACCATCTATCAAGATTCCTGAAGGTCTG 444
Qy 373 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 392
Db 445 AATCTCAGATTAAGTCTGGAGAGACAGTCGCTTGGTCTCAATGGCAGTGGGAAG 504
Qy 393 SerThrValValGlnLeuLeuGlnArgLeuTyArgProAspAspGlyPheIleMetVal 412
Db 505 AGTACGGTAGTCAGCTTCTGCAGAGGTTATATGATCCGATCCGATGATGGCTTATCATGGTG 564
Qy 413 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyArgAspHisIle-GlyValVa 432
Db 565 GATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTTGGAGTGGT 624
```


QY	432	1serGlnGluPro--ValLeuPheGlyThrThrIleSerAsnAsnIleLysTyr-GlyAr	451	313	ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPro	332
Db	625	TAGTCAAGACCTCGTTTTCGGGACACCATCAGTAACATATCAAGTATGGAGC	684	64	GCCTCGCCCAATGCACGAGGAGCGCTTATGAAGTCTTCAAAATATTTGATATATGCCCC	123
QY	451	gAspAspValThrAspGluGlu-MetGluArgAlaAlaArgGluAla-AsnAlaTyrAsp	470	333	SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu	352
Db	685	AGATGATGGACTGATGACACAAATGMAAAGCAACAGGGAAACCAAGGCGAATGAA	744	124	AGTATAGACAGCTTCTCAAAAGATGGGCGCAACACGACAGCAACATACAGGAATCTGGAA	183
QY	471	PhelIleMetGlu-PheProAsnLysPheAsnThrLeuVal--GlyGluLysGlyAlaG	489	353	PhelYsAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu	372
Db	745	TTTATCATGGATTTCCCAAAAAATTTATCATCTGGAAAGGGGAAAAAAGACCTTCA	804	184	TTTAAGATATTCACCTTCAGTTACCCATCTCGAAAAGAAGTTTCAGATCTTGAAGGCGCTC	243
QY	489	nMetSerGlyGlyGlnLysGlnArg	497	373	AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys	392
Db	805	AGGAATGGAAGGCCCAAAACCCCAAG	829	244	AATCTGAAGTGAAGAGCGGACAGACGGTGGCCCTGGTTGGCAACACAGTGGCTGGGAAA	303
RESULT 7				393	SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal	412
LOCUS	BF584668	894 bp	mRNA	linear	EST 12-DEC-2000	
DEFINITION	602098406F1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4218385 5',			304	AGCACAACTGTCCAGCTGATGCAAGGGCTTACGACCCCTAGATGGCATGGTCAATC	363
ACCESSION	BF584668			413	AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal	432
VERSION	BF584668.1	GI:11658386		364	GACGGACAGACATCAGAACCATCATGTGAGGTATCTCGGGGAGATCATTTGTGTGGTG	423
KEYWORDS	EST.			433	SerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAsp	452
SOURCE	house mouse.			424	AGTCAGAAACCTGTGCTGTTTCCACACGATCGCCGAGAACATTCGCTATGCGCCGAGAA	483
ORGANISM	Mus musculus			453	AspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIle	472
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			484	GATGTCCACCATGATGAGATTGAGAAAGCTGTCAAGGAAGCCCAATCCCTATGACTTCATC	543
AUTHORS	1 (bases 1 to 894)			473	MetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGly	492
TITLE	NIH-MGC http://mgs.nci.nih.gov/.			544	ATGAACTGCCCCACCAATTGACACCTGTGTGTGAGAGAGGGCGGACGTGAGTGGG	603
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)			493	GlyGlnLysGlnArgIleAlaAlaArgAlaLeuValArgAsnProLysIleLeuIle	512
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgabbs@email.nih.gov Tissue Procurement: Jeffrey E. Green, M.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.llnl.gov Plate: LLAM9798 row: 1 column: 02 High quality sequence stop: 651.			604	GGACAGAAACAGAGAATCGC-ATTGCCGGGCCCTGGTCCGAATCCCAAGTGCCTTC-661	661
FEATURES	source			513	LeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla	531
	1..894	/organism="Mus musculus"		662	GTGGACGAAGCCACCTCAGC-CTGGATACAGAAAGTGAAGCTGGGGTTCAGGCCGCA	717
		/strain="FVB/N"		RESULT 8		
		/db_xref="taxon:10090"		AK020318	Mus musculus adult male epididymis cDNA, RIKEN full-length enriched	
		/clone_lib="NCI CGAP Co24"		LOCUS	library, clone:9230106F14:similar to MULTIDRUG RESISTANCE PROTEIN 3	
		/lab_host="DH10B (T1 phage-resistant)"		DEFINITION	(P-GLYCOPROTEIN 3) (MDRIA), full insert sequence.	
		/notes="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dr. Average insert size 1.6 kb. Constructed by Life Technologies. Note: this is a NCI CGAP Library."		ACCESSION	AK020318.1	GI:12860872
				VERSION	AK020318	
				KEYWORDS	HTC; CAP trapper.	
				SOURCE	Mus musculus (strain: C57BL/6J) adult male epididymis cDNA to mRNA, clone_lib:RIKEN full-length enriched mouse cDNA library	
				ORGANISM	clone:9230106F14.	
					Mus musculus	
BASE COUNT	263 a	203 c	242 g	185 t		
ORIGIN					Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
Alignment Scores:					1	
Pred. No.:	4,19e-81	Length:	894		Carninci, P. and Hayashizaki, Y.	
Score:	750.00	Matches:	156		High-efficiency full-length cDNA cloning	
Percent Similarity:	84.94%	Conservative:	47		Meth. Enzymol. 303, 19-44 (1999)	
Best Local Similarity:	65.27%	Mismatches:	36		99279253	
Query Match:	27.38%	Indels:	4		PUBMED	
DB:	12	Gaps:	0		10349636	
					2	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
US-09-873-409-7 (1-541) x BF584668 (1-894)						
QY	293	valPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGlu	312			
Db	5	GTCTTCTTTTCCTGTTAAT-GGAGCATTCAGTGTGGACAGGCATCTCCAAATATTGAA	63			

REFERENCE AUTHORS

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuaki, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Hayada, A., Yanamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanishi, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system - 384-format sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, L. M., Stauble, J. F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S., and Havashizaki, Y.

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)

21085660
11217851

5 (bases 1 to 872)

Akachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiraoka, T., Horii, F., Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawaj, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numata, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numaoki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J., Saizawa, H., Saito, R., Sakai, K., Sakai, K., Sano, H., Sasaki, D., Schriml, L., Shibata, K., Shibata, Y., Shingawa, A., Shiraki, T., Sobabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toyota, T., Yamamura, T., Yamanaka, I., Yasunishiki, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

TITLE
JOURNAL

Direct Submission
Submitted (18-AUG-2000) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp,
URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for
further details.

COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5', GAGAGAGAAGGATCAAGACGCTTTTTTTTTTNN 3', cDNA was prepared by using trihalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 185.2. Second strand cDNA was prepared with the primer adapter of sequence [5',

GAGAGAGAGATTCTCGAGTTAATTAATATCCCCCCCCCC 3'}. cDNA was cleaved with BamHI and XhoI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI. Host: DH10B.

FEATURES

```

location="Quarantilla"
i. 872
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:9230106F14"
/db_xref="MGD:MGI:1911565"
/db_xref="taxon:10090"
/clone="9230106F14"
/seq="final"

```

cds

001: 2012
/note="data source:SPTR, source key:P21447, evidence:ISS
putative
similar to MULTIDRUG RESISTANCE PROTEIN 3 (P-GLYCOPROTEIN
3) (MDR1A)"

```

3/ (mkria/
/codon_start=1
/protein_id="BAB32070.1"
/db_xref="GI:12860873"
/db_xref="MGD:1924956"
/translation="MANSRRTGLQETNRYQLQEQPKVGNQAVGIEIFRPADN
DIVLTGLIAGINATAPVMSVLGEISDLHNGCLVQRTYKQNCQSQTQEKLN
DIYMLTGYIGAAIALIEGYQYISFWITTAQRITRKQFPHSLAQDISFQFGLS
ICELNMTMDGDKLDCGDGKILPMPFONISGFSIGLVSLISLWKSLVLVLSLSP.
MASCALCSRMITLSLTSKETDAYSKAGAVAEALISQITVTTAFGA"

```

BASE COUNT	262 a	178 c	187 g	245 t
ORIGIN	MASSAEECSKMIISLISKEEDAYSNAAG			

Alignment Scores:

Pred. No.:	1.91e-80	Length:	872
Score:	744.50	Matches:	151
Percent Similarity:	80.3%	Conservative:	217
Best Local Similarity:	70.56%	Mismatches:	27
Query Match:	27.18%	Indels:	15
DB:	11	Gaps:	2

US-09-873-409-7 (1-541) x AK020318 (1-872)

Ov 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

Db 230 ATGACTCTGGGAATATTAGCATCCATGATAAATGGAGCCACCGTTCCTTTAAATGTCCCTG 289

31 ValT auCluCluMotSovLcnLnI auT oSovCluCuI auVaICl nThrvLnThrv 40

QY 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnIleYr 40

Db 290 GTTTTAGGAGAAATAAGTGATCATTTAATTAATGGATGCCCTAGTACAAACTAACAGAACT 349

1. The first step is to identify the problem or question that needs to be addressed. This involves understanding the context and the specific requirements of the task.

Qy 41 SerPhe-----PheArgLeuThr 46
 | | | | |

350 AATCTCAGACGCGTCAATCGAAGATACTATTGTGGACT
 :::
 409

DB 350 AAATATCAGAACGTGTCTCAGACTCAGGAAAGCTGAATGAAGATATCATGTGTTGACT 409

Qy 47 LeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeu 66

[illegible]

Db 410 CTATATTGGAATAGGAGCAGCTGCCCTCATTTTGGCTATGTACAGATTTCCTTC 469

67 Tmrl,eil,ethrA]aB]zAvcG]nThrI,vB,rqT]eArqI,vsG]npBpBpHeH;sserVa] 86

[illegible]

D_b 470 TGGGTCACTGCAGCCGGCAAACCAAGAAATCCGAAACAGTTT'TTTCATTCAATT 529

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

Qy 87 LeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGluLeuAsnThrArgMet 106

530 TTGGCACAGACATCAGCTGGTTTGATGGCAGTGACATCTGTGAACCTTAACACCCGCATG 589

DD 330 T TGGCACCAAGACATCAGCTGGT T TGGATGGCAAGTACATCTCTGTGACCTTATACACCCCCCGGATG 339

Qy 107 Thr---AspIleAspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeupheGln 125

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

Db 590 ACTGGTGACATCAACAACTCTGTGATGGTATTGGAGATAAGATCCCCCTCTGATGTTTCAG 649

126 AspMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLysGlyTyrPrlLysLeuThr 145

[illegible]

RODIAEPDAKTCQLVSRLLTVDOPKPSFKLVISQGLRSCTOVIGSIVLSMLSPZ
 TMLMVAIVPALMGQVTIMSGSLKLVSRQCEQIARATQVADGDLGNVTRVAPAMEKX
 EERYQABLESCECKAEGRGLTALFQGLSNATFCMWGLTALFGGSLVPAQGLKGD
 RMGSLFVAPQVTPKSGSLVLFQGVQVRLSGAGRVFYMALSPVITPGTGVIPKNDI
 RGSITFONVTVYPCRPGLNVLKDTLPGSKIVLVAQVSGGGKTTVALLLERFYDP
 EAGSVDLHDHLRLTNMLWRGQVIGTFSQEPVLFATIMENIRFKGLDASDEVTA
 AREANAHEFTISPPDQVSGVEERTGTTGGQKQLALAIRIKQPTVLIDENALTA
 DAGESVRQEQALDRASAGRTVLVAHRLSTVRAAKHSITVMANGQVCEAGTHELLKKG
 GSGSLVIRROTDLASLTTPAAKCEKDPKSCSKA"

polyA signal

```
/gene="AUCD8"
/notes="putative"
```

polyA site

```
/gene="Abcb8"
```

```
/note="putative"
```

Case	Year	Age	Sex	Occupation	Location	Time of Day	Time of Year	Weather	Activity	Witnesses	Notes
1	1985	75	M	Farmer	Field	10:00 AM	Spring	Sunny	Working	None	Found dead in field
2	1986	78	F	Teacher	Classroom	3:00 PM	Summer	Hot	Teaching	Students	Found dead in classroom
3	1987	72	M	Engineer	Office	9:00 AM	Autumn	Cloudy	Working	Colleagues	Found dead in office
4	1988	70	F	Retired	Home	11:00 PM	Winter	Cold	Sleeping	Neighbors	Found dead in bed
5	1989	74	M	Doctor	Hospital	7:00 AM	Spring	Foggy	Working	Nurses	Found dead in hospital
6	1990	76	F	Writer	Library	2:00 PM	Summer	Hot	Reading	Librarian	Found dead in library
7	1991	73	M	Engineer	Office	10:00 AM	Autumn	Cloudy	Working	Colleagues	Found dead in office
8	1992	71	F	Retired	Home	11:00 PM	Winter	Cold	Sleeping	Neighbors	Found dead in bed
9	1993	75	M	Farmer	Field	10:00 AM	Spring	Sunny	Working	None	Found dead in field
10	1994	77	F	Teacher	Classroom	3:00 PM	Summer	Hot	Teaching	Students	Found dead in classroom
11	1995	74	M	Engineer	Office	9:00 AM	Autumn	Cloudy	Working	Colleagues	Found dead in office
12	1996	72	F	Retired	Home	11:00 PM	Winter	Cold	Sleeping	Neighbors	Found dead in bed
13	1997	76	M	Doctor	Hospital	7:00 AM	Spring	Foggy	Working	Nurses	Found dead in hospital
14	1998	78	F	Writer	Library	2:00 PM	Summer	Hot	Reading	Librarian	Found dead in library
15	1999	75	M	Engineer	Office	10:00 AM	Autumn	Cloudy	Working	Colleagues	Found dead in office
16	2000	73	F	Retired	Home	11:00 PM	Winter	Cold	Sleeping	Neighbors	Found dead in bed
17	2001	77	M	Farmer	Field	10:00 AM	Spring	Sunny	Working	None	Found dead in field
18	2002	79	F	Teacher	Classroom	3:00 PM	Summer	Hot	Teaching	Students	Found dead in classroom
19	2003	76	M	Engineer	Office	9:00 AM	Autumn	Cloudy	Working	Colleagues	Found dead in office
20	2004	74	F	Retired	Home	11:00 PM	Winter	Cold	Sleeping	Neighbors	Found dead in bed
21	2005	78	M	Doctor	Hospital	7:00 AM	Spring	Foggy	Working	Nurses	Found dead in hospital
22	2006	80	F	Writer	Library	2:00 PM	Summer	Hot	Reading	Librarian	Found dead in library
23	2007	77	M	Engineer	Office	10:00 AM	Autumn	Cloudy	Working	Colleagues	Found dead in office
24	2008	75	F	Retired	Home	11:00 PM	Winter	Cold	Sleeping	Neighbors	Found dead in bed
25	2009	79	M	Farmer	Field	10:00 AM	Spring	Sunny	Working	None	Found dead in field
26	2010	81	F	Teacher	Classroom	3:00 PM	Summer	Hot	Teaching	Students	Found dead in classroom
27	2011	78	M	Engineer	Office	9:00 AM	Autumn	Cloudy	Working	Colleagues	Found dead in office
28	2012	76	F	Retired	Home	11:00 PM	Winter	Cold	Sleeping	Neighbors	Found dead in bed
29	2013	80	M	Doctor	Hospital	7:00 AM	Spring	Foggy	Working	Nurses	Found dead in hospital
30	2014	82	F	Writer	Library	2:00 PM	Summer	Hot	Reading	Librarian	Found dead in library
31	2015	79	M	Engineer	Office	10:00 AM	Autumn	Cloudy	Working	Colleagues	Found dead in office
32	2016	77	F	Retired	Home	11:00 PM	Winter	Cold	Sleeping	Neighbors	Found dead in bed
33	2017	81	M	Farmer	Field	10:00 AM	Spring	Sunny	Working	None	Found dead in field
34	2018	83	F	Teacher	Classroom	3:00 PM	Summer	Hot	Teaching	Students	Found dead in classroom
35	2019	80	M	Engineer	Office	9:00 AM	Autumn	Cloudy	Working	Colleagues	Found dead in office
36	2020	78	F	Retired	Home	11:00 PM	Winter	Cold	Sleeping	Neighbors	Found dead in bed
37	2021	82	M	Doctor	Hospital	7:00 AM	Spring	Foggy	Working	Nurses	Found dead in hospital
38	2022	84	F	Writer	Library	2:00 PM	Summer	Hot	Reading	Librarian	Found dead in library
39	2023										

Alignment Scores:

Pred. No.:	7e-73	Length:	2676
Score:	689.50	Matches:	195
Percent Similarity:	52.42%	Conservative:	98
Best Local Similarity:	34.88%	Mismatches:	187
Query Match:	25.17%	Indels:	79
DB:	11	Gaps:	20

US-09-873-409-7 (1-541) x AK014319 (1-2676)

Qy	1	Met	Ile	Leu	Gly	Ile	Leu	Ala	Ser	Leu	Val	Asn	Gly	Ala	Cys	Leu	Pro	Leu	Met	Pro	Leu	20
471	ATTTGTGTCGTCCTT	AGGTCGGG	CAC	TAGT	GAAT	-----	GTG	CAGAT	CCCGTTG	518												

Qy 21 ValLeuGlyGluMetSerAspAsnLeuIleSer-----GlyCysLeu 34

Db 519 CTCTGGGGCAGCTGGTGAGATTTCGCCAAGTACACGAGGGACCACATGGGAGTTTC 578

Qy 35 valGlnThrAsnThrTyrSerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyVal 54

Db 579 GTGTCAGTCCCGTAAGCTCAGCGTCCAGCTGCTCTACTGTAC-----GGTGTT 629

Qy 55 AlaAlaLeuLeu---PheGlyTyrLeuGlnIleSerLeuTyrPilelleThrAlaAlaArg 73
|||||
56 AlaAlaLeuLeu---PheGlyTyrLeuGlnIleSerLeuTyrPilelleThrAlaAlaArg 73
|||||
57 AlaAlaLeuLeu---PheGlyTyrLeuGlnIleSerLeuTyrPilelleThrAlaAlaArg 73
|||||
58 AlaAlaLeuLeu---PheGlyTyrLeuGlnIleSerLeuTyrPilelleThrAlaAlaArg 73
|||||
59 AlaAlaLeuLeu---PheGlyTyrLeuGlnIleSerLeuTyrPilelleThrAlaAlaArg 73
|||||
60 AlaAlaLeuLeu---PheGlyTyrLeuGlnIleSerLeuTyrPilelleThrAlaAlaArg 73
|||||
61 AlaAlaLeuLeu---PheGlyTyrLeuGlnIleSerLeuTyrPilelleThrAlaAlaArg 73
|||||
62 AlaAlaLeuLeu---PheGlyTyrLeuGlnIleSerLeuTyrPilelleThrAlaAlaArg 73
|||||
63 AlaAlaLeuLeu---PheGlyTyrLeuGlnIleSerLeuTyrPilelleThrAlaAlaArg 73
|||||
64 AlaAlaLeuLeu---PheGlyTyrLeuGlnIleSerLeuTyrPilelleThrAlaAlaArg 73
|||||
65 AlaAlaLeuLeu---PheGlyTyrLeuGlnIleSerLeuTyrPilelleThrAlaAlaArg 73
|||||
66 AlaAlaLeuLeu---PheGlyTyrLeuGlnIleSerLeuTyrPilelleThrAlaAlaArg 73
|||||
67 AlaAlaLeuLeu---PheGlyTyrLeuGlnIleSerLeuTyrPilelleThrAlaAlaArg 73
|||||
68 AlaAlaLeuLeu---PheGlyTyrLeuGlnIleSerLeuTyrPilelleThrAlaAlaArg 73
|||||
69 AlaAlaLeuLeu---PheGlyTyrLeuGlnIleSerLeuTyrPilelleThrAlaAlaArg 73
|||||
70 AlaAlaLeuLeu---PheGlyTyrLeuGlnIleSerLeuTyrPilelleThrAlaAlaArg 73
|||||
71 AlaAlaLeuLeu---PheGlyTyrLeuGlnIleSerLeuTyrPilelleThrAlaAlaArg 73
|||||
72 AlaAlaLeuLeu---PheGlyTyrLeuGlnIleSerLeuTyrPilelleThrAlaAlaArg 73
|||||

DB	630	CAGGGCACTGCTGACCTTTGGATACCTAGTGTCTGCTGTCTCCACAAATT-----GGTGAAGCGC	688
QV	74	GlnThrLysArqIleArqLysGlnPhePheHisserValLeuAlaGlnAspIleGlyTyr	93

Db 684 ATGGCCATGGACATGGGAAAGCCCTTTTCAGTCCCTGCTCGGCAAGACATTGCTTTC 743

Qy 94 PheAspSerCysAspIleGlyGluLeuAsnThrArgMet---ThrAspIleAspLys--- 111

D_b 744 TTTGATGCCAAAAAGACAGGGCAGCTAGTGTAGTCGCTTGACTACTGATGTGCAAGAATTC 803

	Qy	112	-----IleSerAspGlyIleGlyAspPlysIleAlaLeuLeuPhe	129
	Dβ	804	AAGTCATCTTCAAGGTGGTGATCTCCACGGGAATTG----	839

Qy 125 GlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu---ValIysGlyTrpLys 143

840 ---CGAGCTGCACCCAGGTGATTGGTAGCCTGGTGTCCTGCTATGCTGTCCCTCGC 896

QY 144 LeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAlaCysSer 163

897	CTTACCCGTGCTGGCTGTGTCACACCCGCCCTCATGGGAGTGGGCACCCCTGATGGGC	956
-----	--	-----

	Qy	ArgMetValIleSerLeuthrThrSeryLySGluLeuSerAlaLysFserLYSAclayALaval ::: :	187
Dd		TcAGGCCTCCCGAAAGCTGTCTCGCAGTGTTCAGGACGAGATTGCCAGGGCAACAGGTGA 957 TCAGGCCTCCCGAAAGCTGTCTCGCAGTGTTCAGGACGAGATTGCCAGGGCAACAGGTGA 	101

Qy 184 AlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGlu

Db 1017 GCAGATGAGGCCCTTGGCAATGTTCCGACTGTGCGGCCCTTCGCATGGAGAAGAGGGAG 107

Qy 204 LeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrp 223

Db	1077	GAGGAAGCG	:::	----	TATCAAGCAGAA	1097
Qy	224	LeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAsp		----	AlaLysAsp	243
Db	1098	CTGGAGTCATGCTGCTGT		----	AAAGCAGAGA	1127
Qy	244	PheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhe		----	-----	261
Db	1128	CTGGGC-----AGGGCATCCGCTTGTTCGAAGGGCTCTCCAACATCGCTTTCAACTGT		----	CTCAACTGT	1181
Qy	262	PheMetAsnGlyThrTyrGlyLeuAlaPheThrTyrGlyThrSerLeuIleLeuAsnGly		----	LeuAsnGly	281
Db	1182	ATGGTCTTGGGCACC-----CTGTTCAATGGGGCTCCCTTGTGGCTGGCAG		----	CTGGCTGGCAG	1229
Qy	282	Glu---ProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer		----	ValIleHisSer	300
Db	1230	CAGCTGAAGAGGGGAGACCTCATGCTCTTCTGTGGTCTCCACAGACAGTACAGAGGTCT		----	ACAGAGGTCT	1289
Qy	301	SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGly		----	AlaArgGly	319
Db	1290	ATGCCACGCTCTCTGCTCTGTTTGGTCAG-----GTGGTACGTGGGCTG		----	GTGGTACGTGGGCTG	1334
Qy	320	---AlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSer		----	AspAsnPheSer	338
Db	1335	AGTCGGGAGCCGAGTCTCGAATACATGCGCTGAGCCCTGTCATCCCA-----TTG		----	TTG	1388
Qy	339	ThrAlaGlyTyr-----LysProGluSerIleGluGlyThrValGluPheLysAsnVal		----	LeuPheLysAsnVal	356
Db	1389	ACGGGGGCTACTGCAATCCCAACAGGACATTCGTGGTTCCATCACCTTCCAAATGTC		----	TTCCAAATGTC	1448
Qy	357	SerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIle		----	LeuAsnLeuArgIle	376
Db	1449	ACCTTCAGCTACCCCTGCAGACCTGCTCAATGTGCTCAAGGACTTCACCTGAAGCTG		----	TCACCTGAAGCTG	1508
Qy	377	LysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValVal		----	SerThrValVal	396
Db	1509	CCCTCTGGCAAGATTGCGCTCTTGTGGCCCATGCTGGGGAGGAAAGACACCATGTGCC		----	ACACCATGTGCC	1568
Qy	397	GlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAsp		----	ValAspGluAsnAsp	416
Db	1569	TCCTTGTGGAAGCCTCTATGACCCCTGAAGCTGGCTGGTGACGTTGGATGGGCATGAC		----	TGGATGGGCATGAC	1628
Qy	417	IleArgAlaLeuAsnValAlaArgHisTyrArgAspHis---IleGlyValValSerGlnGlu		----	ValValSerGlnGlu	435
Db	1629	CTGGCAACTCTCAACCCCTCGGCTCCGGGGCCAGGTCATAGGTTTCATCAGCCAGGAG		----	TTCATCAGCCAGGAG	1688
Qy	436	ProValLeuPheGlyThrIleSerAsnAsnIleLysTyrGlyArgAspValThr		----	ArgAspValThr	455
Db	1689	CCAGTCTCTGTTTGCACAACACCATCATGAGAGATATCCGATTTTGGGAAGCTGATGCTTCC		----	TGATGCTTCC	1748
Qy	456	AspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPhe		----	IleMetGluPhe	475
Db	1749	GATGAAGAGGTGTACAGCTGCACGAGAAGCCAATGCCACGAGTTTCATCAGCAGCTTC		----	TTCATCAGCAGCTTC	1808
Qy	476	ProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLys		----	SerGlyGlyGlnLys	495
Db	1809	CCGATGGCTACAGCACTGTGGTTGTGAGCGGGGCACAACTTGTCTGTGGTGGCCAGAG		----	TGTGGTGGCCAGAG	1866
Qy	496	GlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGlu		----	IleLeuIleLeuAspGlu	515
Db	1869	CAGCGCTACCAATCGCAGTGGCTTCATCAGCAGCCACAGCTGCTGATCCTGGCAGG		----	TGATCCTGGCAGG	1928
Qy	516	AlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaLeuGluLys		----	AlaLeuGluLys	534
Db	1929	GCACCAAGTCGCTAGATCGAATCCGAGAGGGTGTGACAGGAGCCCTGACCGG		----	CTGACCGG	1985

RESULT 10

AL520322

LOCUS 220300

DEFINITION

5
4
4
4
4
4
4
2
2
2

ACCESSION

AL520322 943 bp mRNA linear EST 13-FEB-2001
AL520322 LTI_NFL004_NBC2 Homo sapiens cDNA clone CSDB006YC15 5
prime, mRNA sequence.
AL520322

5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 260232-265223, 340488-345479, and 484488-489479."

BASE COUNT	115 a	81 c	85 g	124 t
ORIGIN				
Alignment Scores:				
Pred. No.:	7.44e-70			Length: 405
Score:	654.00			Matches: 128
Percent Similarity:	100.00%			Conservative: 1
Best Local Similarity:	99.22%			Mismatches: 0
Query Match:	23.88%			Indels: 0
DB:	9			Gaps: 0

US-09-873-409-7 (1-541) x AA243820 (1-405)

Qy	291	LeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaValProHis	310
Db	18	CTCTCTGTTTCTTTTAGTGTAAATCCATAGCAGTATTGCAATGGAGCAGCAGTCCCTCAC	77
Qy	311	PheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLys	330
Db	78	TTTGAACACCTTCGCAATAGCCCGAGAGCTGCTTCATATTTTCAGGTATTGTATAAG	137
Qy	331	LysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThr	350
Db	138	AAACCCAGTATAGATAACTTTTCCACAGCTGGATATAAACCTGAATCCATAGNAGAACT	197
Qy	351	ValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLys	370
Db	198	GTGGAAATTTAAAAATGTTCTTTCAATTATCCATCAAGACCATCATCAAGATTCTTGAAA	257
Qy	371	GlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySer	390
Db	258	GGTCTGAATCTCAGAAATTAAGTCTGGAGAGACAGTGCCTTGCTGGTCTCAATGGCAGT	317
Qy	391	GlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIle	410
Db	318	GGGNAGAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCGGATGATGGCTTTATC	377
Qy	411	MetValAspGluAsnAspIleArgAla	419
Db	378	ATGGTGGATCAGAAATGACATCAGAGCT	404

RESULT 12	
BJ075336	
LOCUS	
DEFINITION	BJ075336 NIBB Mochii normalized Xenopus tailbud library Xenopus laevis cDNA clone XL057d10 5', mRNA sequence.
ACCESSION	BJ075336
VERSION	BJ075336.1 GI:17520252
KEYWORDS	EST.
SOURCE	African clawed frog. <i>Xenopus laevis</i>
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae; Xenopodinae; <i>Xenopus</i> . 1 (bases 1 to 550)
REFERENCE	Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and Kohara Y.

```

/clone="XL057d10"
/clone_lib="NIBB Mochii normalized Xenopus tailbud
library"
/tissue type="whole embryo"
/dev stage="stage 25"
/notes="vector; pBSRN3; Site 1: NotI; Site 2: EcoRI; cDNAs
were oligo-dT primed and directionally cloned. Staging
according to Nieuwkoop and Faber. Library is subtracted
and was constructed by N. Garrett and A.M. Zorn,
(Wellcome/CRC Institute)."
162 a 109 c 144 g 135 t

```

BASE COUNT	162 a	109 c	144 g	135 t
ORIGIN				

Alignment Scores:		
Pred. No.:	3,818-69	Length:
Score:	650.00	Matches:
Percent Similarity:	85.79%	Conservative:
Best Local Similarity:	66.13%	Mismatches:
Query Match:	23.73%	Indels:
DB:	13	Gaps:
		550
		121

US-09-873-409-7 (1-541) x BJ075336 (1-550)

Qy	349	GlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIle	368
Db	2	GGCGATATAGGAATTCAAAGCATTCATTTTTGGCTACCCATGCCGTCTGGTGTGCAAAAT	61
Qy	369	LeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsn	388
Db	62	CTCTCGGACTGAACCTGAAAGTGGAGCTGGTAAACCATCGCTTTGGTGTGTAGAT	121
Qy	389	GlySerGlyLysSerThrValValGlnLeuGlnArgLeuTyrAspProAspAspGly	408
Db	122	GGTTGGTGTAAAGACACTACCATACAGTTGCTGCAGAGGTTCTATGATCATCATGGA	181
Qy	409	PheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHis	428
Db	182	GAGGTACTTTGGATGGACACGATATCCGTTCCACTAAATGTAATAATGGTTAGAGGAAAT	241
Qy	429	IleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnIleLys	448
Db	242	ATTGGCGTAGTGAGTCAGGAGCCAGTGCTCTTTGGCACCAACCATTCGACAAAATATATCCGC	301
Qy	449	TyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAla	468
Db	302	TACCGAGGGGAATCAGTAAACAGACAAACGAAATGGAACACGCGCAAAAGACCAATGCC	361
Qy	469	TyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAla	488
Db	362	TTTGATTTTATATCCAGGCTTCCTGTATAATTTTAACACGATGGTGGGAAAGAGGTGGC	421
Qy	489	GlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnPro	508
Db	422	CAGTTGAGTGGGGGTGCAAGACGCGGATAGCATTTGCACGCGCCCTGGTACGGAACCT	481
Qy	509	LysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaVal	528
Db	482	AAGATTCTACTTCTTGATGAAGCCACATCAGCTCTGGATACACAAAAGTGAAGCTATTGTA	541
Qy	529	GlnAlaAla 531	
Db	542	CAAGCTGCC 550	
RESULT 13			
BG298756			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
BG298756 803 bp mRNA linear EST 21-FEB-95 602396681F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:4511157 mRNA sequence. BG298756.1 GI:13063728 EST. house mouse. Mus musculus			

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 803)
 NIH-MGC <http://mgs.nci.nih.gov/>
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-r@mail.nih.gov
 Tissue Procurement: The Cepko Laboratory
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Plate: LLAM10394 row: b column: 22
 High quality sequence stop: 698.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

Location/Qualifiers
 1. .803
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="IMAGE:4511157"
 /tissue_type="retina"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 3.3 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH MGC Library."
 240 a 162 c 205 g 196 t

FEATURES
 source

BASE COUNT
 ORIGIN

Alignment Scores:
 Pred. No.: 1,25e-68 Length: 803
 Score: 648.00 Matches: 149
 Percent Similarity: 68.99% Conservative: 49
 Best Local Similarity: 51.92% Mismatches: 55
 Query Match: 23.66% Indels: 37
 DB: 12 Gaps: 3

US-09-873-409-7 (1-541) x BG298756 (1-803)

Qy 168 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluVal 187
 Db 15 TCATTACTGATAGGAACCTCCATGCTTATGCAAAAGCTGGAGAGTCTGGAAGATC 74
 Qy 188 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 207
 Db 75 TTACGACCATCAGAACTGTGATTCGTTGGA-CGACAAAAGGAAGAACTTGAA----- 127
 Qy 208 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 227
 Db 127 ----- 127
 Qy 228 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 247
 Db 128 -----AGTACATATACAACTTGAAGAAGCTAAAGCGTGGGATTAAG 172
 Qy 248 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 267
 Db 173 AAGCTATACGGCCCAACATCTCATGGTGCAGCTTTCTCTTATCATGATCATAT 232
 Qy 268 GlyLeuAlaPheTrpTyrGlyThrSerLeuLeuLeuAsnGlyGluProGlyTyrThrIle 287
 Db 233 GCTCTGGCAATCTCGTATGGGACTTCCTTGGTCATCTCCAAAGAA-----TACTCTAT- 285
 Qy 288 GlyThrValLeuAlaValPhePheSerValIleHisSerTyrCysIleGlyAlaAla 307
 Db 286 GGACAAAGTGCCTACTGCTCTCTTCCGTGTAAATGAGCAATTCATGTTGGACGGCA 345
 Qy 308 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 327
 Db 308 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 327

Db 346 TCTCCAATATTGAAGCCTTCGCCATGCAGGAGCAGCTTATGAAGTCTTCAAAATA 405
 Qy 328 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 347
 Db 406 ATTGATAATAAGCCAGTATAGACAGCTTCTCAAGAGTGGGCACAAACAGACAATA 455
 Qy 348 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 367
 Db 466 CAAGGAATCTGGAATTTAAGATATTCACTTCACTTCCATCTCGAAAAGAAGTTGAG 525
 Qy 368 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 387
 Db 526 ATCTTCAGAGGCGCTCAATCTGAAGTGAAGAGCGGACAGCGTGGCCCTGTGGTGAAC 585
 Qy 388 AsnGlySerGlyLysSerThrValValGlnLeuGlnArgLeuTyrAspProAspAsp 407
 Db 586 AGTGGCTGTGGAAGAAAGCACAACCTGTCCAGCTGATCAAGGCTCTACGACCCTTA-GAT 644
 Qy 408 GlyPheIleMetValAspGluAsnAspIleArgAlaLeu-AsnValArgHis-TyrArg- 426
 Db 645 GGCATGTCAGTATCCAGCGACGACATCAGAACCATCCATGTGAGGTATCTGGAGGG 704
 Qy 427 AspHisIleGlyValValSerGln-GluProValLeuPheGlyThrThrIleSerAsnAs 446
 Db 705 GATCAT---GGGTGTGTGAGTCAAGGAAGCACTGGCTTTGTTGCCAACCCAGATCGGAGAA 761
 Qy 446 nleLysTyrGly 450
 Db 762 AATTCGTTATGGG 774

RESULT 14

LOCUS

DEFINITION
 EST489713 MHAM Medicago truncatula/Glomus versiforme mixed EST
 library cDNA clone pMHAM-51L19 5' end, mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1. .780
 /organism="Medicago truncatula/Glomus versiforme mixed EST
 library"
 /cultivar="Medicago truncatula genotype A17"
 /db_xref="taxon:119092"
 /clone="pMHAM-51L19"
 /tissue_type="roots"
 /dev_stage="Roots harvested at 10, 17, 22, 31 and 38 days
 post-inoculation with Glomus versiforme. The library was
 made from a mixture of RNA from each of these stages."
 /lab_host="E. coli strain XL0R"
 /note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
 XhoI; cDNA was prepared from polyA+ enriched RNA from
 roots harvested at 10, 17, 22, 31 and 38 days

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 31, 2003, 02:24:20 ; Search time 79.2267 Seconds
(without alignments)
2094.145 Million cell updates/sec

Title: us-09-873-409-7

Perfect score: 2739

Sequence: 1 MILGTILSLVNGACILPLMLP.....SESKSAVQAALKEKDTFRYSF 541

Scoring table: BLOSUM62

Ygapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing First 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-O=/cgn2_1/USPTO/spool/US09873409/runat_27032003/115422_19259/app_query_fasta_1.7544
-DB=Issued_Patents NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09873409 @CGN 1 1 578 @runat_27032003/115422_19259 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents NA:
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1552.5	56.7	4233	3	US-09-120-513-1
2	1552.5	56.7	4233	4	US-09-450-105-1
3	1552.5	56.7	4669	2	US-08-752-447-1
4	1552.5	56.7	4669	4	US-09-316-167-1
5	1546.5	56.5	4646	1	US-08-181-471-2
6	1546.5	56.5	6505	2	US-08-793-610-5
7	1546.5	56.5	9318	2	US-08-793-610-6
8	1545	56.4	4264	2	US-08-784-649A-1
9	1545	56.4	4264	2	US-08-784-649A-5
10	1544.5	56.4	4669	6	5206352-3
11	1525.5	55.7	4669	2	US-08-583-276-18
12	960	35.0	4224	1	US-08-612-521-1

13	935.5	34.2	4047	2	US-08-612-734B-1	Sequence 1, Appli
14	916	33.4	4800	2	US-08-612-734B-3	Sequence 3, Appli
15	914	33.4	4002	2	US-08-996-545-1	Sequence 1, Appli
16	914	33.4	4002	2	US-08-996-545-3	Sequence 3, Appli
17	914	33.4	4002	4	US-09-328-320-1	Sequence 1, Appli
18	914	33.4	4002	4	US-09-328-320-3	Sequence 3, Appli
19	898	32.8	2726	1	US-08-461-823-1	Sequence 1, Appli
20	897.5	32.8	3924	1	US-08-395-246C-1	Sequence 1, Appli
21	849.5	31.0	6143	1	US-08-612-521-3	Sequence 3, Appli
22	785.5	28.7	3792	4	US-09-351-224E-10	Sequence 10, Appli
23	747.5	27.3	3924	2	US-08-996-644-3	Sequence 3, Appli
24	747.5	27.3	3924	3	US-09-352-552-3	Sequence 3, Appli
25	747.5	27.3	3927	2	US-08-996-644-1	Sequence 1, Appli
26	747.5	27.3	3927	3	US-09-352-552-1	Sequence 1, Appli
27	688.5	25.1	3999	4	US-09-351-224E-9	Sequence 9, Appli
28	646	23.6	2376	1	US-08-394-880B-1	Sequence 1, Appli
29	626.5	22.9	2061	4	US-09-061-764A-17	Sequence 17, Appli
30	626	22.9	3909	1	US-08-232-537-1	Sequence 1, Appli
31	621.5	22.7	1959	4	US-09-061-764A-4	Sequence 4, Appli
32	574.5	21.0	2244	4	US-09-061-764A-18	Sequence 18, Appli
33	571.5	20.9	1749	4	US-09-134-001C-1893	Sequence 2774, Ap
34	529	19.3	1743	4	US-09-134-001C-2774	Sequence 63, Appli
35	519.5	19.0	7760	4	US-08-961-527-63	Sequence 6, Appli
36	501.5	18.3	5120	3	US-08-772-270A-6	Sequence 1, Appli
37	501.5	18.3	8370	2	US-08-488-706-1	Sequence 1, Appli
38	491	17.9	4411529	4	US-09-103-840A-1	Sequence 1, Appli
39	489	17.9	13188	4	US-08-961-527-70	Sequence 70, Appli
40	489	17.9	4403765	4	US-09-103-840A-2	Sequence 2, Appli
41	483	17.6	7721	3	US-08-772-270A-14	Sequence 14, Appli
42	482	17.6	28804	2	US-08-592-874-1	Sequence 1, Appli
43	482	17.6	28804	3	US-09-096-942-2	Sequence 2, Appli
44	482	17.6	28804	3	US-09-096-867-2	Sequence 2, Appli
45	460	16.8	6492	4	US-08-961-527-188	Sequence 188, App

ALIGNMENTS

RESULT 1

US-09-120-513-1
; Sequence 1, Application US/09120513
; Patent No. 6025160
; GENERAL INFORMATION:
; APPLICANT: Brun, Kimberly
; APPLICANT: Chenery, Richard
; APPLICANT: Ellens, Harma
; APPLICANT: Field, John
; APPLICANT: Yue, Lin
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE
; TITLE OF INVENTION: SEQUENCES ENCODING RAT MDR1B2 AND
; TITLE OF INVENTION: SCREENING METHODS THEREOF
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY:
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/120,513
; FILING DATE: 22-JUL-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: King, William T
; REGISTRATION NUMBER: 30,954
; REFERENCE/DOCKET NUMBER: GPO008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5015

TELEFAX: 610-270-5090
 TELEEX:
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 4233 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: Genomic DNA
 US-09-120-513-1

Alignment Scores:

Pred. No.: 1.69e-172 Length: 4233
 Score: 1552.50 Matches: 308
 Percent Similarity: 72.20% Conservative: 92
 Best Local Similarity: 55.60% Mismatches: 105
 Query Match: 56.68% Indels: 49
 DB: 3 Gaps: 5

US-09-873-409-7 (1-541) x US-09-120-513-1 (1-4233)

QY 1 MetileuGlylleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
 DB 170 ATGGCTCTGGAACTCTCGCTGCTATCATCCAGCAACCCCTGCTTCCCTCCTGATGCTG 229
 QY 21 ValLeuGlyGluMetSerAspAsnLeuLeu----- 30
 DB 230 GTGTTCGATACATGACAGATAGTTTACCCACGACAGACCCGCATCTGCCGAGGCTT 289
 QY 31 -----SerGlyCysLeuValGlnThrAsnThrTyrSerPhePheArgLeu----- 45
 DB 290 ACTAATCAAGTGAATCAACAGTACACAGACCGCTCAGCAGACAGCAGCTGGAGGAGAC 349
 QY 46 -----ThrLeuTyrTyrValGlylleGlyValAlaLeuLeuPheGlyTyr 61
 DB 350 ATGGCCATGTACGCTACTATTACACGGGCATTGCTGCCGTGTGCTCATGCTTGCCTAC 409
 QY 62 ileGlnileSerLeuTyrPheIleThrAlaAlaArgGlnThrLysArgIleArgGln 81
 DB 410 ATCCAGGTTTCACCTTGGTCTGGAGCTGGGAGACAAATACACAGATTTAGGCAGAG 469
 QY 82 PhePheHisSerValLeuAlaGlnAspileGlyTrpPheAspSerCysAspileGlyGlu 101
 DB 470 TTTTTCATGCATCATGAATCAGGAGATAGCTGGTTTGACGTGAATGACGCTGGGAG 529
 QY 102 LeuAsnThrArgMetThr---AspileAspLysIleSerAspGlylleGlyAspLysIle 120
 DB 530 CTCAACACCCGGCTCACAGATGACGCTCTCCAAATTAATGACGGAATGGTGACAACTT 589
 QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
 DB 590 GGAATGTTCTTCAGTCCATCAACGACATTTTACGCCGGTTTAAATAGATTTATAGT 649
 QY 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
 DB 650 GGTGGAGCTTAACCTTGTAAATTTGGCCGTCAGCCCTCTTATTGGGTGTCATCTGCC 709
 QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
 DB 710 ATGTGGGCAAGGTAAGTACTGACTTCTATTCTAATAGGAACCTCCAGGCTTATGCCAAGCT 769
 QY 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
 DB 770 GGACAGTGTCCGAGAGAGTCTTAGCAGCCATCAGACTGTGATGGTTGGAGGACAA 829
 QY 201 GlyLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
 DB 830 AAGAAGGAACCTTGA----- 844
 QY 221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
 DB 845 -----AGGTACAATAAAATTTAGAGAA 868

QY 241 AlaLysAspPheGlylleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
 DB 869 GCTAAAGAGATTGGCATAAAGAAAGCCATCAGCGCAACATTTCCATAGTATTGCTAC 928
 QY 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsn 280
 DB 929 CTGTTGGTCTATGCTGCTTATGCACTGGCATTCGTGTATGGGACCTCTTGGTCTCTCA 988
 QY 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
 DB 989 AATGAA-----TATTCTATGGACAAGTCTTACCGTCTTCTCTCTATTTATTGGGG 1042
 QY 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
 DB 1043 ACITTCAGTATTGGACATTTAGCCCCCAACATAGAGCCTTTGCAAAATGCAAGAGGGCA 1102
 QY 321 AlapheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
 DB 1103 GCCTATGAAATCTTCAAGATAATTGATAATGAGCCCAAGCATTCACACGCTTCTCAACCAAG 1162
 QY 341 GlyTyrLysProGluSerIleGlyThrValGluPheLysAsnValSerPheAsnTyr 360
 DB 1163 GGACAAACACACAGATATATGGGAAATTTGGAATTTTAAATGTTTACTTCAACTAC 1222
 QY 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
 DB 1223 CCATCAGAAAGTGAAGTAAAGATCTTGAAGGGCCTCAACCTGAAGGTGAAGAGCGGGCAG 1282
 QY 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
 DB 1283 ACGGTACGCCCTGTTGGCAACAGTGGCTGTGGGAAAGCAACACTGTCTCCAGCTGCTGCAG 1342
 QY 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspileArgAlaLeu 420
 DB 1343 AGGCTCTACACCCCATAGAGGGCGAGGTGAGTATGACGGACAGGACATCAGACCATC 1402
 QY 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
 DB 1403 AATGTGAGGTATCTCGGGAAATCAATTTGGGTGTGAGTCAGGAACCCGTGCTGTTGCC 1462
 QY 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
 DB 1463 ACCACGATTCGCCAAACATTCGCTATGGCCGAGAAACGTCACCATGGATGAGATAGAG 1522
 QY 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
 DB 1523 AAGCTGTCAAGGAAGCCATGCTATGACTTATATGAAGTGCCTCCCAAAATTTTAAAC 1582
 QY 481 ThrLeuValGlyLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
 DB 1583 ACCCTGTTGTCAGAGAGGGGCGCAGCTGAGTGGGGGACAGAAACAGAGGATCGCCATT 1642
 QY 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
 DB 1643 GCCCGGGCCCTGTCGCAACCCCAAGATCTTTGTTGGATGAGGCGCACGTCAGCCTTG 1702
 QY 521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
 DB 1703 GACACAAAGCGAAGCCGCTGTTTCAGGCCCTCTCGATTAAG 1744

RESULT 2

US-09-450-105-1
 ; Sequence 1, Application US/09450105
 ; Patent No. 6169166
 ; GENERAL INFORMATION:
 ; APPLICANT: Kimberly Anne Brun
 ; APPLICANT: Richard James Chenery
 ; APPLICANT: Harma Ellens
 ; APPLICANT: John Anthony Feild
 ; APPLICANT: Lin Yue
 ; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES
 ; TITLE OF INVENTION: ENCODING RAT MDR1B2 AND SCREENING METHODS THEREOF
 ; FILE REFERENCE: GP-50008-D1

; CURRENT APPLICATION NUMBER: US/09/450,105
; CURRENT FILING DATE: 1999-11-29
; EARLIER APPLICATION NUMBER: 09/120,513
; EARLIER FILING DATE: 1998-07-22
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 4233
; TYPE: DNA
; ORGANISM: HOMO SAPIENS
US-09-450-105-1

Alignment Scores:

Pred. No.: 1.69e-172 Length: 4233
Score: 1552.50 Matches: 308
Percent Similarity: 72.20% Conservative: 92
Best Local Similarity: 55.60% Mismatches: 105
Query Match: 56.68% Indels: 49
DB: 4 Gaps: 5

US-09-873-409-7 (1-541) x US-09-450-105-1 (1-4233)

QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
DB 170 ATGGCTCTGGGAACCTCTCGCTGCTATCATCCACGGAACCTCTCCCTCTGTGCTG 229
QY 21 ValLeuGlyGluMetSerAspAsnLeuIle----- 30
DB 230 GTGTTCGGATACATGACATAGTTTATCCCAAGCAGACCGCGATCTGCGGAGCGTT 289
QY 31 -----SerGlyCysLeuValGlnThrAsnThrTyrSerPheArgLeu----- 45
DB 290 ACTAATCAAAGTGAATCAACAGTACACAGCCGTGACGACAGCAGTCTGGAGGAGGAC 349
QY 46 -----ThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
DB 350 ATGGCCATGTACCGCTTACATTAACAGCGGATTTGGTGGCGGTGTCTATCGTTGCTAC 409
QY 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
DB 410 ATCCAGGTTTCACTTTGGTCCCTGGCAGCTGGGAGACAAATACACAAGATTAGGCAGAG 469
QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
DB 470 TTTTTCATGCCATCATGAATCAGGAGATAGGCTGTTGACCTGAATGACGCTGGGAG 529
QY 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
DB 530 CTCACACCCGGCTCAGATGACGATGACGCTCTCCAAAATTAATGACGGAATTTGTTGACAACTT 589
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
DB 590 GGAATGTCTTTTCAGTCCATACACATTTTACCGCGTTTATATAGATTTATAGT 649
QY 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
DB 650 GGTGGAGCTACCTCTGTAATTTGGCGTCCAGCCCTCTTATTTGGTGTGTCATCTGCC 709
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGlnLeuSerAlaTyrSerLysAla 180
DB 710 ATGTGGCAAGGTACTGACTTCTATTACTTAATAAGGAACCTCCAGGCTTATGCGAAGCT 769
QY 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
DB 770 GGAGCAGTGTCCGAAGAAGTCTTAGCAGCCATCAGAACTGTGATTTGGTGGAGACAA 829
QY 201 GluLysGlnLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
DB 830 AAGAAGGAACCTGAA----- 844
QY 221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
DB 845 -----AGGTACAATAAAATTTAGAAGAA 868

QY 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
DB 869 GCTAAAGAGTTGGCATAAAGAAAGCCATCAGCGCAACATTTCCATAGTATTCCTAC 928
QY 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsn 280
DB 929 CTGTTGGTCTATGCGTCTATGCACTGGCATCTGTGGTATGGAGACCTCTTGTCTCTCA 988
QY 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
DB 989 AATGAA-----TATTCTATTGACAAAGTGTCTACCGTCTTCTCTCTATTATTATGGGG 1042
QY 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
DB 1043 ACTTTCAGTATTGGACATTTAGCCCCAACATAGAGCCCTTTGCAAAATCAAGAGGGCA 1102
QY 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
DB 1103 GCCTATGAAATCTTCAAGATAATTGATATGAGCCCAAGCATTTGACAGCTTCTCAACAAG 1162
QY 341 GlyTyrLysProGluSerIleGlyThrValGluPheLysAsnValSerPheAsnTyr 360
DB 1163 GCACACAAACACAGACAGTATATGGAATTTGGAATTTAAATAATGTTACTTCAACTAC 1222
QY 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
DB 1223 CCATCAGGAAGTGAATTAAGATCTTGAAGGGCTCAACCTGAAGGTGAAGAGCGGCAG 1282
QY 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
DB 1283 ACGGTAGCCCTGTTGGCAACAGTGGCTGTGGAAAGCACAACTGTCCAGCTGTGCGAG 1342
QY 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
DB 1343 AGGCTCTACACACCCATAGAGCGCGAGGTTCAGTATCGACGGACAGACATCAGGACCATC 1402
QY 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
DB 1403 AATGTGAGGTATCTCGGGAATCATTTGGGTGGTGGTTCAGGAACCCGCTGTCTTGTGCC 1462
QY 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
DB 1463 ACCACGATTGGCGAAACATTCGCTATGCGCGAGAAAGCTCACCATTGATGATAGATAG 1522
QY 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
DB 1523 AAAGCTGTCAAGAGCCCAATGCTATGATTCATCATGAACTGCCCCACAAATTTAAC 1582
QY 481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlnLysGlnArgIleAlaIle 500
DB 1583 ACCCTGTTGGTGGAGAGGGCGCGAGTGTGGGGGACAGAAACAGAGGATCGCCATT 1642
QY 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
DB 1643 GCCCGGCGCTGCTGCGCAACCCCAAGATCTTTTGTGGATGAGGCGCACGTCAGCCTTG 1702
QY 521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
DB 1703 GACACAGAAGCGAAGCCGTGTTTACGGCGCTCTCGGATAAG 1744

RESULT 3

US-08-752-447-1

; Sequence 1, Application US/08752447

; Patent No. 5994088

; GENERAL INFORMATION:

; APPLICANT: Mechneter, Eugene

; APPLICANT: Roninson, Igor B

; TITLE OF INVENTION: Methods and Reagents for Preparing and

; TITLE OF INVENTION: Using Immunological Agents Specific for p-glycoprotein

; NUMBER OF SEQUENCES: 2

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: McDonnell Boenhen Hulbert & Berghoff Ltd.

STREET: 300 South Wacker Drive, Seventh Floor
 CITY: Chicago
 STATE: Illinois
 COUNTRY: USA
 ZIP: 60606
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 FILING DATE: 15-NOV-1996
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: No. 599408nan, Kevin E
 REGISTRATION NUMBER: 35,303
 REFERENCE/DOCKET NUMBER: 95,1121
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 312-913-0001
 TELEFAX: 312-913-9808
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 4669 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: 5'UTR
 LOCATION: 1..424
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 425..4264
 FEATURE:
 NAME/KEY: 3'UTR
 LOCATION: 4265..4669
 US-08-752-447-1

Alignment Scores:
 Pred. No.: 1.99e-172 Length: 4669
 Score: 1552.50 Matches: 305
 Percent Similarity: 73.29% Conservative: 101
 Best Local Similarity: 55.05% Mismatches: 99
 Query Match: 56.68% Indels: 49
 DB: 2 Gaps: 6

US-09-873-409-7 (1-541) x US-08-752-447-1 (1-4669)

Qy 1 MetileLeuGlyIleLeuAlaSerLeuValAenGlyAlaCysLeuProLeuMetProLeu 20
 Db 575 ATGGTGGGAACTTGGCTGCCATCATCCATGGGCTGCATCTCTCATGATGCTG 634
 Qy 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
 Db 635 GTGTTTGGAGAAATGACAGATATCTTTGCAAAATGACGAAATTTAGAGATCTCATGTCA 694
 Qy 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
 Db 695 AACATCAATAAGAGATGATCAATGATACAGGGTTCTTTCATGAATCTGGAGAAAGAC 754
 Qy 44 -----ArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
 Db 755 ATGACCGATGATGCTATTATTCAGTGGAAATGGCTGGGGTGGCTGGTGGCTGTAC 814
 Qy 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgGln 81
 Db 815 ATTCAGGTTTCATTTGGTGCTGCGACGTGCGAAGACAAATACACAAAATTTAGAAAACAG 874
 Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
 Db 875 TTTTTCATGCTATTATCGCAGAGATAGGCTGGTTTGTATGTCACGATGTTGGGGAG 934

Qy 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
 Db 935 CTTTAAACCCGACTTACAGATGATGCTCCAGATTAAATGAAGGAATTTGGTGACAAAATT 994
 Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
 Db 995 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGGATTTACACGT 1054
 Qy 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
 Db 1055 GGTGGAAGCTAACCTTGGATTTGGCCATGATGCTCTGTTCTTGGACTGTGAGTGTCT 1114
 Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaIleValLysAla 180
 Db 1115 GTCTGGCAAGATACCTATCTTCTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 1174
 Qy 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
 Db 1175 GGAGCAGTAGCTGAAGAGGTCTTTGGCAGCAATTAGAACTGTGATTGCATTGGAGGCAA 1234
 Qy 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
 Db 1235 AAGAAAGAACTTGAA----- 1249
 Qy 221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
 Db 1250 -----AGGTACAACAACAAAATTTAGAAAGAA 1273
 Qy 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
 Db 1274 GCTAAAAGAAATTTGGGATAAAGAAAGCTATTACAGCCCAATATTTCTATAGTGTGCTTTC 1333
 Qy 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
 Db 1334 CTGCTGATCTATGCATCTTAAGTCTGGCCCTCTGGTATGGACCACTTGGTCTCTCA 1393
 Qy 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
 Db 1394 GGGAA-----TATCTATTGGCAAGTACTCACTGTTATTTCTTCTATTAAATGGG 1447
 Qy 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
 Db 1448 GCTTTTAGTGTGGCAGGCACTCTCCAAGCATTTGCAAGCATTTGCAATGCAAGAGGAGCA 1507
 Qy 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
 Db 1508 GCTTATGAATCTTCAAGATAATTGATAAGCCAGTATTGACAGCTATTCGAAGAGT 1567
 Qy 341 GlyTyrLysProGluSerIleGlyThrValGluPheLysAsnValSerPheAsnTyr 360
 Db 1568 GGGCACAACACAGATAATATAAGGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 1627
 Qy 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
 Db 1628 CCATCTCGAAAAGAAAGTTAAGATCTTGAAGGGCTGAAACCTGAAGGTGCGAGAGTGGCAG 1687
 Qy 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
 Db 1688 ACGTGGCCCTGGTTGGAAACAGTGGCTGTGGAGAGCAACACAGTCCAGCTGATGCAG 1747
 Qy 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
 Db 1748 AGGCTCTATGACCCACAGAGGGGATGGTCAGTGTTCATGACAGGATATTAGGACCATA 1807
 Qy 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
 Db 1808 AATGTAAAGTTTCTACGGGAAATCAATTTGGTGTGGTGTAGTCAGGAACCTGATTGTTGCC 1867
 Qy 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
 Db 1868 ACCACGATAGCTGAAACATTTGCTATGGCCGTGAAATGTCACCATGGATGAGATTGAG 1927
 Qy 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480


```

Db 1928 AAAGCTGTCAAGGAAGCCATGCTATGATCTTATCATGAACCTGCTCATAAATTGAC 1987
Qy 481 ThrLeuValGlyGlyGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
Db 1988 ACCCTGTTGGAGAGAGAGGGCCAGTGGTGGTGGGCGAGAGGATGCGCAATT 2047
Qy 501 AlaArgAlaLeuValArgAnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
Db 2048 GCACGTGCCCTGTTGCAACCCCAAGATCCTCTGCTGGATGAGGCCACGTGAGCCTTG 2107
Qy 521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
Db 2108 GACACAGAAAGCGAAGCAGTGGTTCAGGTGGCTCTGGATAAG 2149

RESULT 4
US-09-316-167-1
; Sequence 1, Application US/09316167
; Patent No. 6365357
; GENERAL INFORMATION:
; APPLICANT: Mechneter, Eugene
; APPLICANT: Roninson, Igor B
; TITLE OF INVENTION: Methods and Reagents for Preparing and
; TITLE OF INVENTION: Using Immunoligal Agents Specific for P-glycoprotein
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff Ltd.
; STREET: 300 South Wacker Drive, Seventh Floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/316,167
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/752,447
; FILING DATE: 15-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: NO. 6365357han, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 95,1121
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-9808
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4669 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: 1..424
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 425..4264
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: 4265..4669
US-09-316-167-1

Alignment Scores:
Pred. No.: 1,99e-172 Length: 4669
Score: 1552.50 Matches: 305
Percent Similarity: 73.29% Conservative: 101
```

```

Best Local Similarity: 55.05% Mismatches: 99
Query Match: 56.68% Indels: 49
DB: 4 Gaps: 6

US-09-873-409-7 (1-541) x US-09-316-167-1 (1-4669)

Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAenGlyAlaCysLeuProLeuMetProLeu 20
Db 575 ATGGTGGTGGGAACCTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 634
Qy 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
Db 635 GTGTTGGAGAAATGACATATCTTTGCAAAATGCAGGAAATTTAGAGATCTGATGTCA 694
Qy 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
Db 695 AACATCACTAATAGAAGTATATCAATGATACAGGGTCTTTCATGAATCTGGAGGAAGAC 754
Qy 44 -----ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db 755 ATGACCAGGTATGCTTATTTATACAGTGGAAATTTGGTGGCTGGGGTGTCTGCTGCTTAC 814
Qy 62 IleGlnIleSerLeuTyrIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db 815 ATTCAGTTTTCATTTTGGTGGCTGGCAGCTGGAGCAAAATACACAAATTTAGAAACAG 874
Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101
Db 875 TTTTTCATGCTATTATGCGACAGGAGATAGGCTGGTTTGTGATGTCACCATGTTGGGGAG 934
Qy 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db 935 CTTAACACCCGACTTACAGATGATGCTCCAGATTAATGAAGGAATTTGGTGACAAAT 994
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db 995 GGAATGTTCTTTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGGATTTACAC 1054
Qy 141 GlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db 1055 GGTGGAGCTAACCTTGTGATTTTGGCCATCAGTCTCTTCTTGGACTGTCAGGTGCT 1114
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db 1115 GTCTGGGCAAGATATACTATCTTCTATTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 1174
Qy 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db 1175 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACCTGTGATTCGATTTGGAGGACAA 1234
Qy 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
Db 1235 AAGAAGAAGACTTGAA----- 1249
Qy 221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
Db 1250 -----AGGTACAACAAAAATTTAGAAGAA 1273
Qy 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
Db 1274 GCTAAAGAAATTTGGGATAAAGAAAGCTATTACAGCAATTTTCTATAGGTGCTGCTTTC 1333
Qy 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrGlyThrSerLeuIleLeuAsn 280
Db 1334 CTGCTGATCTATGATCTTATGCTGGGGCTTCTGGTATGGGACACCTTGGTCTCTCA 1393
Qy 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
Db 1394 GGGGA-----TATTCATTGGACAAGTACTCAGTGTATTCTTTCTGATTAAATGGG 1447
Qy 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
Db 1448 GCTTTTAGTGTGGACAGGCATCTCCAGCATTTGAAGCATTTTGCAAAATTCAGAGGAGCA 1507
```



```
Db 1175 GGAGCAGTACTGAAGAGGCTTGGCAGCAATTAGAACTGTGATTCATTTGGAGACAA 1234
Qy 201 GLuysGluLeuGlnArgSerPheLeuAasnIleThrArgTyrAlaTrpPheTyrPhe 220
Db 1235 AAGAAAGAACTTGAA-----1249
Qy 221 ProGlnTrpLeuSerCysValLeu***PheValArgTyrThrGlnAenLeuLysAsp 240
Db 1250 -----AGGTACACAAAAATTTAGAGAA 1273
Qy 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
Db 1274 GCTAAAGAAATTCGGATAAAGAGCTATTACAGCCAAATTTCTATAGGTGCTGCTTC 1333
Qy 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrPheGlyThrSerLeuIleLeuAen 280
Db 1334 CTGCTGATCATCATCTATGCTCTGGCCCTTGGTATGGACCACCTTGGTCCCTCA 1393
Qy 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
Db 1394 GGGGAA-----TATTCTATTGGCAAGTACTCACTGATTCTTTCTGTATTAAATGGG 1447
Qy 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
Db 1448 GCCTTTAGTGTGGACAGGCATCTCCAAGCATTTCAAGCATTTGCCAATGCCAAGAGGCA 1507
Qy 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAenPheSerThrAla 340
Db 1508 GCTATGAATCTTCAAGATAATTGATAAGCCAGTATTACAGCTATTGCAAGAGT 1567
Qy 341 GlyTyrLysProGluSerIleGlyThrValGluPheLysAenValSerPheAenTyr 360
Db 1568 GGGCAAAACACAGATAATTAAGGGAAATTCGAATTCAGAAATGTTCACTTCAGTTAC 1627
Qy 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAasnIleuArgIleLysSerGlyGlu 380
Db 1628 CCATCTCGAAAGAAAGTAAAGTCTTGAAGGGCTGAACTGAAAGTGCAGAGTGGGCAG 1687
Qy 381 ThrValAlaLeuValGlyLeuAasnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
Db 1688 ACGTGGCCCTGTTGGAAACAGTGGCTGTGGGAAGACACACAGTCCAGCTGATGCGAG 1747
Qy 401 ArgLeuTyrAspProAspArgPheIleMetValAspGluAenAspIleArgAlaLeu 420
Db 1748 AGGCTCTATCACCCACAGAGGGGATGGTCAGTGTTCATGGACAGGATATTAGGACCAT 1807
Qy 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
Db 1808 AATGTAAGGTTTCTACGGGAAATCATTTGGTGTGAGTCAGGAACCTGTATTGTTGCC 1867
Qy 441 ThrThrIleSerAenAenIleLysTyrGlyArgAspValThrAspGluGluMetGlu 460
Db 1868 ACCACAGATAGCTGAAACATTCGCTATGGCCGGAATAATGTCCACCTGATGATGATGAG 1927
Qy 461 ArgAlaAlaArgGluAlaAenAlaTyrAspPheIleMetGluPheProAenLysPheAen 480
Db 1928 AAAGCTGTCAAGGAAGCAATGCCTATGACTTTATCATGAACTGCCTCATAAATTGAC 1987
Qy 481 ThrLeuValGlyGlnLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
Db 1988 ACCCTGGTGGAGAGAGGGGCCAGTGTAGTGTGGGAGAGCAGAGGATCGCCATT 2047
Qy 501 AlaArgAlaLeuValArgAenProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
Db 2048 GCACGTGCCCTGGTTCGCAACCCCAAGATCTCTGCTGGATGAGGCCAGCTCAGCCCTG 2107
Qy 521 AspSerGluSerLysSerAlaValGlnAlaLeuGluLys 534
Db 2108 GACACAGAAAGCGAAGCAGTGGTTTCAGGTGCCTCGATTAAG 2149
```

RESULT 6
US-08-793-610-5

```
; Sequence 5, Application US/08793610
; Patent No. 5858744
; GENERAL INFORMATION:
; APPLICANT: BAUM, Christopher
; APPLICANT: STOCKING-HARBERS, Carol
; APPLICANT: OSTERTAG, Wolfram
; TITLE OF INVENTION: RETROVIRAL VECTOR HYBRIDS AND THE USE THEREOF
; NUMBER OF INVENTIONS: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nikaido, Marmelstein, Murray & Oram LLP
; STREET: 655 Fifteenth Street N.W. Suite 330
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/793,610
; FILING DATE: 07-MAR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DB P 44 31 973.8
; FILING DATE: 08-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DB 195 03 952.1
; FILING DATE: 07-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP95/03175
; FILING DATE: 10-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Berman, Richard J.
; REGISTRATION NUMBER: 39,105
; REFERENCE/DOCKET NUMBER: P1614-7007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)638-5000
; TELEFAX: (202)638-4810
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6505 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA
; US-08-793-610-5

Alignment Scores:
Pred. No.: 1,78e-171 Length: 6505
Score: 1546.50 Matches: 305
Percent Similarity: 73.10% Conservatives: 100
Best Local Similarity: 55.05% Mismatches: 100
Query Match: 56.46% Indels: 49
DB: 2 Gaps: 6

US-09-873-409-7 (1-541) x US-08-793-610-5 (1-6505)

Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAenGlyAlaCysLeuProLeuMetProLeu 20
Db 1967 ATGGTGGTGGGAACCTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 2026
Qy 21 ValLeuGlyGluMetSerAsp-----AenLeuIleSer 31
Db 2027 GTGTTGGGAAGATGACAGATATCTTGGCAATGCAGGAATTTAGAGATCTGATGTC 2086
Qy 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe-----43
Db 2087 AACATCACTAATAGAGTATGATATACAGGGTTCCTTCATGAATCTGGAGGAAGAC 2146
Qy 44 -----ArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db 44 -----ArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
```


MOLECULE TYPE: DNA
US-08-793-610-6

Alignment Scores:

Pred. No.: 3,256-171 Length: 9318
Score: 1546.50 Matches: 305
Percent Similarity: 73.10% Conservative: 100
Best Local Similarity: 55.05% Mismatches: 100
Query Match: 56.46% Indels: 49
DB: 2 Gaps: 6

US-09-873-409-7 (1-541) x US-08-793-610-6 (1-9318)

```
QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
DB 1926 ATGGTGGTGGAACTTTGGCTGCCATCATCCATGGGGTGGACTTCCTCTCATGATGCTG 1985
QY 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
DB 1986 GTGTTTGGAGAAATGACAGATATCTTTGCCAAATGCGAGAAATTTAGAAAGATCTGATGTC 2045
QY 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
DB 2046 AACATCACTAATAGAAAGTATGATATCATAGGCTTCTTCATGATCTGGAGGAGAC 2105
QY 44 -----ArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
DB 2106 ATGACACGAGTATGCTTATTTATACAGTGAATTTGGTGGGTGCTGGTGTGCTTAC 2165
QY 62 IleGlnIleSerLeuTrpIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
DB 2166 ATTCAGGTTTCTTTGGTGGCTGGCAGCTGGAGACAAATACACAAATTTAGAAACAG 2225
QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
DB 2226 TTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTGATGTCACGATGTTGGGAG 2285
QY 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
DB 2286 CTTAACACCGACTTACAGATGATGCTCTAAGATTAATGAGATTTGGTGACAAAT 2345
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
DB 2346 GGAATGTTCTTTTCAGTCAATGCGAACATTTTTCACCTGGGTTTATAGTAGGATTTACACGT 2405
QY 141 GlyTriPlysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
DB 2406 GGTGGAAGCTAACCCCTTGATTTTGGCCATCAGTCCCTGTTTGGACTGTCAGCTGCT 2465
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
DB 2466 GTCTGGCAAGATATCTATCTTACTATGATTAAGAACTCTTAGGATGCAAAAGCT 2525
QY 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
DB 2526 GGAGCAGTAGCTCAAGAGGCTTTGGCAGCAATTAGAATCTGATGCTGATTTGGAGACAA 2585
QY 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
DB 2586 AAGAAAGAACTTGAA----- 2600
QY 221 ProGlnTrpLeuLeuSerCysValLeu**PheValArgTyrThrGlnAsnLeuLysAsp 240
DB 2601 -----AGGTACAAACAAAATTTAGAGAA 2624
QY 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
DB 2625 GCTAAAGAAATTTGGATAAGAAAGCTATTACAGCCCAATTTCTATAGTGTGCTTTC 2684
QY 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
DB 2685 CTGCTGATCTATGATCTTATGCTCTGGCCTCTCTGGTATGGGACCACTTGGTCTCTCA 2744
```

```
QY 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
DB 2745 GGGGAA-----TATTTCTATGGCAAGACTACTCAGTATTTCTTTCTATTAATTTGGG 2798
QY 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
DB 2799 GCTTTTAGTGTGGACAGCATCTCCAGCATTTGAAGCATTTTGCAAATGCAAGAGGCA 2858
QY 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
DB 2859 GCTTATGAATCTTCAAGATAATTGATAATAAGCAAGATTATGACAGCTATTTCGAAGAGT 2918
QY 341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
DB 2919 GGGCAAAACCAAGATAATAATTAAAGGAATTTTGAATTCAGAAATGTTCACTTCAGTTAC 2978
QY 361 ProSerArgProSerIleLysLysLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
DB 2979 CCATCTCGAAAGAGATTGAAGATCTTGAAGGCTTGAACCTGAAGGTGCAGGTGGGCGAG 3038
QY 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
DB 3039 ACGGTGGCTGTTGGAAACAGTGGCTGTGGGAAGAGCACACAGCTCCAGCTGATGCAG 3098
QY 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
DB 3099 AGGCTCTATGACCCACAGAGGGGATGGTTCAGTGTGTGATGGACAGGATATTAGGACATA 3158
QY 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
DB 3159 AATGTAAGTGTCTTACGGGAATCATTTGTTGTTGATGGACAGGAACTGTATTGTTTGGC 3218
QY 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluMetGlu 460
DB 3219 ACCAGTAGCTGAAACATTTGCTATGCGCGTGAATGTCAACATGATGATGATGAG 3278
QY 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
DB 3279 AAGCTGTCAAGAGCAATGCTATGATCTTATCATGAAACTGCTCATAAATTTGAC 3338
QY 481 ThrLeuValGlyLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
DB 3339 ACCCTGGTTGGAGAGAGGGGCCAGTTGAGTGGTGGCAGAGAGGATCGCCATT 3398
QY 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
DB 3399 GCACGTGCTGCTGGTTCGCAACCCCAAGATCTCTCTGCTGATGAGGCCAGCTCAGCTTG 3458
QY 521 AspSerGluSerLysSerAlaValGlnAlaLeuGluLys 534
DB 3459 GACACAGAAAGCAAGCAGTGGTTCAGGTGGCTCTGGATAAG 3500
```

RESULT 8

US-08-784-649A-1
; Sequence 1, Application US/08784649A
; Patent No. 5830697
; GENERAL INFORMATION:
; APPLICANT: Sikic, Branimir I
; APPLICANT: Chen, Gang
; TITLE OF INVENTION: P-GLYCOPROTEIN MUTANT RESISTANT TO
; TITLE OF INVENTION: CYCLOSPORIN MODULATION
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 2200 Sand Hill Road
; CITY: Menlo Park
; STATE: CA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/784,649A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sherwood, Pamela J
REGISTRATION NUMBER: Reg.No. 5830697 36,677
REFERENCE/DOCKET NUMBER: 06037/007001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-322-5070
TELEFAX: 415-854-0875
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4264 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-784-649A-1

Alignment Scores:
Pred. No.: 1,31e-171 Length: 4264
Score: 1545.00 Matches: 306
Percent Similarity: 73.29% Conservative: 100
Best Local Similarity: 55.23% Mismatches: 98
Query Match: 56.41% Indels: 50
DB: 2 Gaps: 7

US-09-873-409-7 (1-541) x US-08-784-649A-1 (1-4264)

```
QY 1 MetileLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
DB 292 ATGGTGGGGAACCTTGGCTGCCATCATCCATGGGCTGGACTTCCTCTCATGATCGT 351
QY 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
DB 352 GTGTTTGGAGAAATGACAGATATCTTTGCCAAATGCGAGAAATTTAGAGATCTGATGCA 411
QY 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
DB 412 AACATCTAATAAGAGTATATCAATGATACAGGGTCTTTCATGAATCTGGAGGAAGAC 471
QY 44 -----ArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
DB 472 ATGACCCAGGTATGCTATTATACAGTGGAAATGGTCTGGGGTCTGGTTCGTCTTAC 531
QY 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
DB 532 ATTCAGGTTTCATTTGGTGGCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 591
QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
DB 592 TTTTTCATGCTATATATCGCAGAGATAGGCTGGTTGATGTCACGATGTTGGGGAG 651
QY 102 LeuAsnThrArgMetThr-----AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
DB 652 CTTAACACCCGACTTACATGATGCTCTCCAGAAATTAATGAGGAATTTGGTACAAATTT 711
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
DB 712 GGA---ATGTTTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGGATTTTACAGT 768
QY 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
DB 769 GGTGGGAAGCTTAACCTTGTGATTTTGGCCATCATGCTCTGTTCTGGACGTGCTGCTGCT 828
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
DB 829 GTCTGGGCAAGAGACTACTATCTTCTACTGATAAAGACTCTTAGCGTATGCAAAAGCT 888
QY 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
```

```
DB 889 GGAGCAGTAGCTGACAGAGCTCTTGGCAGCAATTAGAACCTGTGATTCATTTGGAGACAA 948
QY 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
DB 949 AAGAAAGACTTTGAA----- 963
QY 221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
DB 964 -----AGGTACACACAAAATTTAGAGAA 987
QY 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
DB 988 GCTAAAAGAAATTTGGGATAAGAAAGCTATTACAGCAATATTTCTATAGTGTCTGCTTTC 1047
QY 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
DB 1048 CTGCTGATCTATGATCTTATGCTCTGGCTTCTGGTATGGGACCACTTGGTCTCTCA 1107
QY 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
DB 1108 GGGGAA-----TATCTATTGGACAAGTACTCACTGTATTCTTTCTGTATTAAATGGG 1161
QY 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
DB 1162 GCTTTTAGTGTGGACAGCATCTCCAGCATTTGAAAGCATTTGCAAAATGCAAGAGAGCA 1221
QY 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
DB 1222 GCTTATGAATCTTCAAGATATTGATATAGCCAAATTTGACAGCTATTGCAAGAGT 1281
QY 341 GlyTyrLysProGluSerIleGlyThrValGluPheLysAsnValSerPheAsnTyr 360
DB 1282 GGGCACAAACCATTAATATTAGGGAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 1341
QY 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
DB 1342 CCATCTCGAAAAGAAATTAAGATCTTTGAAGGGCTCAACTGAAGGTGACAGATGGGCGC 1401
QY 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
DB 1402 ACGTGGCTCTGTTGGAAACAGTGGCTGGGAAGACACACACAGCTCCAGCTGATGCAG 1461
QY 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
DB 1462 AGGCTCTATGACCCACAGAGGGGATGGTCAGTGTGTGATGGACAGCATATTAGCACATA 1521
QY 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
DB 1522 AATGTAAGTGTCTACGGGAAATCATTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGC 1581
QY 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspValThrAspGluMetGlu 460
DB 1582 ACCAGTATGCTGAAACCATTCGCTATGGCCGTAATAATGTCACCATGATGATGATGAG 1641
QY 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
DB 1642 AAAGCTGTCAGAGCAACCAATGCTTATGATCTTATCATGAACCTGCTCATAAATTTGAC 1701
QY 481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
DB 1702 ACCCTGTTGGACAGAGAGGGGCCAGTGTGAGTGGTGGGAGAGAGAGATTCGCCATT 1761
QY 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
DB 1762 GCACGTGCCCTGTTTGGAAACCCCAAGATCTCTCTGCTGGATGAGGCCAGCTCAGCCTTG 1821
QY 521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
DB 1822 GACACAGAAAGCAAGCAGTGGTTTCAAGTGGCTCTCGATAAG 1863
```

RESULT 9
US-08-784-649A-5
; Sequence 5, Application US/08784649A

Patent No. 5830697
 GENERAL INFORMATION:
 APPLICANT: Sikic, Branimir I
 APPLICANT: Chen, Gang
 TITLE OF INVENTION: P-GLYCOPROTEIN MUTANT RESISTANT TO
 TITLE OF INVENTION: CYCLOSPORIN MODULATION
 NUMBER OF SEQUENCES: 5
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fish & Richardson
 STREET: 2200 Sand Hill Road
 CITY: Menlo Park
 STATE: CA
 COUNTRY: USA
 ZIP: 94025
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/784,649A
 FILING DATE:
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Sherwood, Pamela J
 REGISTRATION NUMBER: Reg. No. 5830697 36,677
 REFERENCE/DOCKET NUMBER: 06037/007001
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-322-5070
 TELEFAX: 415-854-0875
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 4264 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 US-08-784-649A-5

Alignment Scores:
 Pred. No.: 1,31e-171 Length: 4264
 Score: 1545.00 Matches: 306
 Percent Similarity: 73.29% Conservative: 100
 Best Local Similarity: 55.23% Mismatches: 98
 Query Match: 56.41% Indels: 50
 DB: 2 Gaps: 7

US-09-873-409-7 (1-541) x US-08-784-649A-5 (1-4264)

Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
 Db 292 ATGGTGGTGGAACTTTGGCTGCCATCCATCGGGCTGACCTCTCTCATGATGCTG 351
 Qy 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
 Db 352 GTGTTGGGAAATGACAGATATCTTTGCAAAATGCAGGAAATTTAGAGATCTGATGTC 411
 Qy 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
 Db 412 AACATCACTAATAGAGTGTATCATGATACAGGGTCTTCATGATCTGGAGGAGAC 471
 Qy 44 -----ArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
 Db 472 ATGACCAAGGTATGCCATTATTACAGTGAATTTGGTGGGTGCTGGTGGTCTGCTTAC 531
 Qy 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
 Db 532 ATTCAGGTTTCATTTGGTGGCTGGCAGCTGGAAGACAAATACACAAATTTAGAAACAG 591
 Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
 Db 592 TTTTTCATGCTATAATGCGACAGGAGATAGGCTGTTGATGTCACCGATGTTGGGAG 651

Qy 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
 Db 652 CTTAACACCCGACTTACAGATGATGTCTCAAGATTAATGAAGGAATGGTGCAAAAT 711
 Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
 Db 712 GGA--ATGTTCCAGTCATGCGACATTTTTCACCTGGGTTTATAGTAGGATTTACACGT 768
 Qy 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
 Db 769 GGTGGAGCTAAACCTTGTGATTTTGGCCATCAGCTCTGTTCTTGGACTGTGACGTGCT 828
 Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
 Db 829 GTCTGGGCAAGATACTATCTTCATTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 888
 Qy 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
 Db 889 GGAGCAGTAGCTGAGAGGTCTTGGCAGCAATTAGACCTGTGATTCATTTGGAGGACAA 948
 Qy 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
 Db 949 AAGAAAGAACTTGAA----- 963
 Qy 221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
 Db 964 -----AGGTACAAACAAATAATTTAGAGAA 987
 Qy 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
 Db 988 GCTAAAGAAATTTGGATAAAGAGCTATTACAGCCAAATTTCTATAGTGTGCTTTC 1047
 Qy 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
 Db 1048 CTGCTGATCTATGATCTTATGCTGCTGCTTCTGATGGACCACTTGTGCTCTCA 1107
 Qy 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
 Db 1108 GGGGA-----TATTCTATTGGACAAAGTACTCCTGTTATTTCTGTTAATTTGGG 1161
 Qy 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyVala 320
 Db 1162 GCTTTTAGTGTGGACAGCATCTCCAGCATTGAAGCAATTTGCAATGCAAGAGGACGA 1221
 Qy 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
 Db 1222 GCTTATGAATCTTCAAGATAATTGATATAAGCCAAGTATTGACAGCTATTGCAAGAGT 1281
 Qy 341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
 Db 1282 GGGCACAAACCAAGATAATTAAGGGAAATTTGGAATTCAGAAATGTTCACTTTCAGTTAC 1341
 Qy 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
 Db 1342 CCATCTCGAAAAGAGTTAAGATCTTGAAGGGCTTGAACCTGAAGGTGCGAGTGGGCAG 1401
 Qy 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValGlnLeuLeuGln 400
 Db 1402 ACGTGGCCCTGGTTGGAACACAGTGGCTGGGAAGAGCACACACAGTCCAGCTGATGACG 1461
 Qy 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
 Db 1462 AGGCTCTATGACCCACAGAGGGGATGGTCAGTGTGTGATGGCAGCAGGATATTAGGACATA 1521
 Qy 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
 Db 1522 AATGTAAGGTTTCTACGGGAATCATTTGTTGTTGAGTCAGGAACCTGTTATTGTTGCC 1581
 Qy 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspValThrAspGluGluMetGlu 460
 Db 1582 ACCACAGTAGCTGAAACATTTGCTATGGCCGTGAAATGTCAACCATGATGATGATGAG 1641
 Qy 461 ArgAlaAlaArgGluAlaAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480

Qy	32	GlyCysLeuValGlnThr-----AsnThrTyrSerPhe-----	43
Db	695	AAATCACTAATAGAGTGATATCAATATACAGGGTCTTCATGAAATCGAGGAAGAC	754
Qy	44	-----ArgLeuThrLeuTyrValGlyLeGlyValAlaLeuLePheGlyTyr	61
Db	755	ATGACACAGGTATGCCATTATTTACATAGTGAATTGGTGTGGGGTGTGGTTCGCTTAC	814
Qy	62	IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln	81
Db	815	ATTTCAGGTTCATTTTGGTGCTGGCAGCTGGAGACAAATATACACAAATATGAAAAACAG	874

RESULT 11

Qy	82	PhosphoHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyCglu	101
Db	875	TTTTTTCATGCTATAATGACGACAGAGATAGCTGGTTTGATGTGCACGATGTGGCG	934
Qy	102	LeuAsnThrArgMetThr---AspIleAspIlySileSerAspGlyIleGlyAspIlySile	120
Db	935	CTTTAACACCCGACTTACAGATGATGCTCTAAGATTAATGAACCTATTGTGTGACAAAATT	994
Qy	121	AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuVallys	140
Db	995	GGAAATGTTCTTTCAGTCAATGCGCAACATTTTTCATCTGGGTTTATAGTAGGATTTACACGT	1054
Qy	141	GlyTTrpIysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla	160
Db	1055	GGTTGGNAGCTAACCCCTGTGATTGTGGCCATCAGTCTCTGTCTTGGAGCTGCAGCTGCT	1114
Qy	161	AlaCysSerArgMetValIleSerLeuThrSerIysGluLeuSerAlaTyrSerIlyAla	180
Db	1115	GTCTGGCGAAGATCATCTCTTCATTACTGATAAAGAACTCTTAGCGGTATGCAAAAGCT	1174

	181	GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln	200
Qy	181	GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln	200
	1175	GGAGCAGTAGCTCAAGAGGTCTTGGCAGCAATTAGAACTGTGATTCATTTGGAGGCACAA	1234
Db	1175	GGAGCAGTAGCTCAAGAGGTCTTGGCAGCAATTAGAACTGTGATTCATTTGGAGGCACAA	1234
	201	GluIysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyArgAlaTrpPheTyPhe	220
Qy	201	GluIysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyArgAlaTrpPheTyPhe	220
	1235	AAAGAAGAACTTGAA-----	1249
Db	1235	AAAGAAGAACTTGAA-----	1249

Qy	221	ProGlnTrpLeuSerCysValLeu***pheVala	gTyrThrGlnAsnLeuLysasp	240
			::: :::	
Db	1250	-----	---AGGTACAACAAAATTTAGAGAA	1273
Ov	241	AlaLysAspPheGlyIleLysArqThrIleAlaSerLysValSerLeuGlyValaValTyr		260

QY 241 *Almya snasp* *incell* |||CTCT|||GCTGTTTCTATAGGCTGCTTTC 1333

Qy 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsn 280

Db 1334 CTGCTGATCTATGCATCTTATGCTCTGGCCCTTCTGGTATGGACACCTTGGTCTCTCA 1393

Qy 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPheSerValIleHisSer 300
 | | | | | | | | | | | | | | | | |
 .. .

Db 1394 GGGGAA-----TATTCTATTGGACAAGTACTCACTGTATTCTTTTCTGTATTATTTGG 1447

Qy 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320

Db 1448 GCTTTTAGTGTGGACAGGCATCTCCAAGCATTGAAGCATTTGCAATGCAAGAGGCA 1507

Qy 321 AlaPheHisIlePheGlnValIleAaspLysProSerIleAaspAsnPheSerThrAla 340

Db 1508 GCTTATGAAATCTTCAAGATAATTGATAATAAAGCCCAAGTATTGACAGTGATTTCGAAGAGT 1567

Qy 341 GlyTyrLysProGluSerIleGluGlyThrValGluPhelYsAsnValSerPheAsnTyr 360
Db 1558 GCGCCACAAACCCGACATATATTAATGGGGAATTTCAGAAAATCTTTCACCTCAGTTAC 1627

DB 1568 GGGCACAACCAGATAAATTAAGGGAATTCAGAAAGTTCACCTTCAGTAAAC

Ov 361 ProSerArgProSerIlelvsvllleuLvsvGlvleuAsnleuArgillelvsvSerGlvglu 380

QY	361	ProserArgProSerIleLysIleLeuValGlyLeuAsnMetuGlyCysSerGlyPhe	1687
Dβ	1628	CCATCTCGAAAAGAAGATTAAAGATCTTGAAGGCCCTGAACCTGAAGGTGCAGGTGGGCAG	1687

Qy 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400

```
Db 1688 ACAGTCCCTGGTGGAAACAGTGGTGGGAAGACGCAACAGTCCAGCTGATGCAG 1747
Qy 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
Db 1748 AGGCTCTATGACCCACAGAGGGGATGGTCAGTGTTCATGGACAGGATATTAGACCATA 1807
Qy 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
Db 1808 AATGTAAGGTTTCTACGGGAATCATTTGGTGTGTGAGTCAGAACCTGTATTGTTGCC 1867
Qy 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
Db 1868 ACCACGATAGCTGAACAATTCGTATGGCGGTGAAATGTCACCATGGATGAGATTGAG 1927
Qy 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
Db 1928 AAAGCTCTCAAGGAAGCAATGCTATGACTTTATCATGAATGCTCTCAATAATTGAC 1987
Qy 481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
Db 1988 ACCCTGGTTGGAGAGAGAGGCCCCAGTTGAGTGGTGGCAGAGCAGAGGATCGCCATT 2047
Qy 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
Db 2048 GCAGCTGCCCTGGTTCGAAACCCCAAGATCTCTGCTGGATGAGGCCAGCTCAGCCTTG 2107
Qy 521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuLys 534
Db 2108 GACAGAAAGCGAAGCAGTGGTTTCAGGTGGCTCTGGATAAG 2149
```

RESULT 12

```
US-08-612-521-1
; Sequence 1, Application US/08612521
; Patent No. 5786463
; GENERAL INFORMATION:
; APPLICANT: Peery, Robert B
; APPLICANT: Skatrud, Paul L
; APPLICANT: Thornevell, Susan J
; TITLE OF INVENTION: MULTIPLE DRUG RESISTANCE GENE OF
; TITLE OF INVENTION: CRYPTOCOCCUS NEOFORMANS
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: USA
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,521
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hamilton, Amy E
; REGISTRATION NUMBER: 33,894
; REFERENCE/DOCKET NUMBER: X-9693
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3169
; TELEFAX: 317-276-1294
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4224 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
```

```

; LOCATION: 1..4224
US-08-612-521-1
Alignment Scores:
Pred. No.: 1,47e-102 Length: 4224
Score: 960.00 Matches: 210
Percent Similarity: 56.89% Conservative: 112
Best Local Similarity: 37.10% Mismatches: 182
Query Match: 35.05% Indels: 62
DB: 1 Gaps: 8

US-09-873-409-7 (1-541) x US-08-612-521-1 (1-4224)
Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 466 ATGGTCTTGGCGCTTGTGCTTGGCGTGGCAGCGGATCTTCCCAACCTTTAATGACTTTA 525
Qy 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
Db 526 ATATTCCGGTCGATTAAACACAGCTTTTCACGAATTATGCTGTCAATT---GCGAACCAATA 582
Qy 41 SerPhePheArgLeuThr-----LeuTyrTyrValGlyIleGlyValAlaAlaLeuIle 46
Db 583 TCCCAAGGGGAGCTTACTCCCGAGAGCTCTGGCGCTTTTACAGCAGCCCAAGATGATCTC 642
Qy 47 -----LeuTyrTyrValGlyIleGlyValAlaAlaLeuIle 58
Db 643 AAGACTCAATCTGGCCACAAATGCCCTTTATCTGATGCCATTTGGCATTTGGAATGTTCTA 702
Qy 59 PheGlyTyrIleGlnIleSerLeuThrIleIleThrAlaAlaArgGlnThrLysArgIle 78
Db 703 GCGACTTGGCTTTACATGTTTCATCTGGAATGTTTACTTGGCGAGCTCAACTCTAAAGGATT 762
Qy 79 ArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAsp 98
Db 763 AGAGAGCGTTACTTGGCTGCAGTGGCTTTAGACAGGAGATTGCCCTACTTTGATGATCTGGGC 822
Qy 99 IleGlyGluLeuAsnThrArgMet---ThrAspIleAspLysIleSerAspGlyIleGly 117
Db 823 GCGGAGAAAGTCCCACTCGCATTCAGACCGATTGTCCCTTCCCAAGAGGGAACATCT 882
Qy 118 AspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGly 137
Db 883 GAAAAGGTCGCTCTCGTATTCCAATATGCTGGTACTTTTGTCTGCGGTTTGTGTAGCT 942
Qy 138 LeuValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAla 157
Db 943 TTTGTCCGGTCACCTCGTCTTGGCGGCGCTCTGTGTTCCATCTCTGCAGTCATCATGCTC 1002
Qy 158 SerAlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyr 177
Db 1003 TGGCGCGGTATTATGATGACGGCCATATGGCCAAATATGGGACTCGGCTCTTGATCACATC 1062
Qy 178 SerLysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPhe 197
Db 1063 GCGAAAAGCGGGCAGCTTGGCAGAAAGAGGTATAGGAAGTATATCAGGACTGTTCAGGCTTT 1122
Qy 198 ArgAlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTip 217
Db 1123 ---GGCAAGGAAAGATCTTGGGTGCAAAATT----- 1152
Qy 218 PheTyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsn 237
Db 1153 -----GCCGATCAC 1161
Qy 238 LeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGly 257
Db 1162 ATTGAGCAGACGCAAGATTGTGCGTAGAAAAGGCTCCACTTTTGAAGCTTTGTTGAGC 1221
Qy 258 AlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeu 277
Db 1222 ATCATGTTCTTCTGTCATACGCCGCTTATGCGCTTGCCTTCTTCTACGGTGCATCTC 1281
```



```
Db 808 CTGGTCATGTTATGGAGGTGGGTGAGGTATTATGTGAAGTACAGCAAGAAATCTATT 867
Qy 175 SerAlaTyrSerLysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrVal 194
Db 868 GAAATTTATGGTCTGCTGGAAGTCTGCGGAAGAGTCACTAGCTCCATCGGAATGCT 927
Qy 195 IleAlaPheArgAlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAenIleThrArg 214
Db 928 ACCGCTTTGGCACTCAGGATAAGTCGCCAAG----- 960
Qy 215 TyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyr 234
Db 961 -----CAATAC 966
Qy 235 ThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysVal 254
Db 967 GAAACCATCTGGCTGAGGTGAAATAATGGGGCGTCAACACAGGTCACTCTTGGTATG 1026
Qy 255 SerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGly 274
Db 1027 ATGATTGGTGTATGTTCCGTATCATGTTCTCGAAGTATGTTCTCGGTTCTCGATGGGA 1086
Qy 275 ThrSerLeuIleLeuAenGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhe 294
Db 1087 TCTCGCTTCGTTGCGTAAGAA-----GTCAACGTGGGCCCAAGTCTGACAGTTTG 1140
Qy 295 PheSerValIleHisSerTyrCysIleGlyAlaValProHisPheGluThrPhe 314
Db 1141 ATGCTATCTGATCGTTGTTGAGTTTGGGCACGTGCCCCCAATGTCAGGCCCTTT 1200
Qy 315 AlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIle 334
Db 1201 ACGAATGGTGTGTCGGCGCGGCGAAGATTACAGCAGGATTGACCGCAGATCGCCACTG 1260
Qy 335 AspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLys 354
Db 1261 GACCCCTATTCTGACGAAGGAAGGTACTACACATTTCGAGGAATATCGAATTTTCG 1320
Qy 355 AsnValSerPheAsnTyrProSerArgProSerIleLysIleLysGlyLeuAsnLeu 374
Db 1321 AATGTCAAACACATCTACCTTTCAAGACCCGAGTTACAGTCATGGAAGATGCTCTTTA 1380
Qy 375 ArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAenGlySerGlyLysSerThr 394
Db 1381 TCGATGCCCGCGGAAGACTACCGCATTGTGGGCCCATCTGCTGCGAAGAGTACT 1440
Qy 395 ValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGlu 414
Db 1441 GTTCTCGCTTGTGGAGCGCTTTTACCTTCCAGTAGGAGGCCAGGTATTGCTGGACGGC 1500
Qy 415 AsnAspIleArgAlaLeuAenValArgHisTyrArgAspHisIleGlyValValSerGln 434
Db 1501 CATGATATCAAAACCCCTCAACCTCCGTGGTGGCGAGCAGATCTCTCTTGTGAGCCAG 1560
Qy 435 GluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly----- 450
Db 1561 GAACCTGTTCTTTTCAGCACACCATCTTTAGAAACATCGACATGCTGCTTATGGCACC 1620
Qy 451 -----ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGlu 465
Db 1621 AAATTCGACATGATGATCGAAGGACGAAGATTAGAGAGCTGTTGAGATGCGGCCAGAA 1680
Qy 466 AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 485
Db 1681 GCCAATGCTCATGATTTTATTATGGCTCTGCTGAAGGTTACGATACGAATGCGGTCAG 1740
Qy 486 LysGlyAlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal 505
Db 1741 CGTGGTTCTTACTTTTCAGGAGGTGAGAAGCAAGTATGCTGCTGCGCCATTGTC 1800
Qy 506 ArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 525
Db 1801 AGTGACCCCAAGATTCTGTTGATGAAGCTACATCAGCTTTGGATACCAAGTCCGAG 1860
```

```
Qy 526 SerAlaValGlnAlaAlaLeuGluLys 534
Db 1861 GCGGTCGTCCTCAAGCGCTCTTGATAAA 1887

RESULT 14
US-08-612-734B-3
; Sequence 3, Application US/08612734B
; Patent No. 5914246
; GENERAL INFORMATION:
; APPLICANT: Peery, Robert B.
; APPLICANT: Skatrud, Paul L.
; APPLICANT: Tobin, Matthew B.
; TITLE OF INVENTION: Multiple Drug Resistance Gene of
; TITLE OF INVENTION: Aspergillus Fumigatus
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center, DC1501
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: U.S.A.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,734B
; FILING DATE: 08-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Craig, Anne I.
; REGISTRATION NUMBER: 32,976
; REFERENCE/DOCKET NUMBER: X-9681
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-354-9570
; TELEFAX: 617-354-4043
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4800 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
US-08-612-734B-3

Alignment Scores:
Pred. No.: 2,86e-97 Length: 4800
Score: 916.00 Matches: 202
Percent Similarity: 55.47% Conservative: 102
Best Local Similarity: 36.86% Mismatches: 182
Query Match: 33.44% Indels: 62
DB: 2 Gaps: 7

US-09-873-409-7 (1-541) x US-08-612-734B-3 (1-4800)
Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 693 ATTCTCTCGGTTTCAGTCAGCTCCAGCGCTTCCAGGGCATATCTCTGGGTATATGCCCTAT 752
Qy 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
Db 753 -----CAT 755
Qy 41 SerPhePhe---ArgLeuThr-----LeuTyrTyrValGlyIleGlyValAlaAla 56
Db 756 GAGTTCATCAAACTGACTAAGAAATGTGCTTTACTTTGTGTATCTCGGTATTTGCCGAG 815
Qy 57 LeuIlePheGlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLys 76
Db 816 TTTGTACAGTCTATGTCACGACCGTGGGTTTCATTTATATTACTGGCGAATCTCACACAG 875
```


QY	77	ArgileArglysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSer	96
DB	876	AAGATCGTGAATAATTATCTTCAGGCTATCTCGAGCAGAAATATGGCTTACTTCGACAAG	935
QY	97	CysAspIleGlyLeuLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGly	115
DB	936	TTGGGCGCGGTGAAGTACACCGGTATCACTGCTGATACCAACCTGATCCAGGACGCC	995
QY	116	IleGlyAspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAla	135
DB	996	ATCTCTGAGAAAGTTGGTCTCTCACTTTGACCGCATTCGCCACATTTGTAAACCGCATTTATT	1055
QY	136	ValGlyLeuValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIle	155
DB	1056	GTCCGCTACGTCAAGTATTGGAAGTTGGCTCTGTATCTGTCTCACTCAACTATCGTCGCGCTG	1115
QY	156	MetAlaSerAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSer	175
DB	1116	GTATCGGTTATGGAGTGGGTGGTCAAGGTTATTATTTGGAAGTACACCAAGAATCTATTGAA	1175
QY	176	AlaTrpSerLysAlaGlyAlaValAlaGluLeuValLeuSerSerIleArgThrValIle	195
DB	1176	AGTTATGGTCTGGTGGAACTGTCCGGAGAGATCATCAGCTCCATTCGGAATGCTACC	1235
QY	196	AlaPheArgAlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyr	215
DB	1236	GCATTTCGGCACTCAGGATAAGCTCGCCAAG-----	1265
QY	216	AlaTrpPheTyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThr	235
DB	1266	-----CHAATACGAA	1274
QY	236	GlnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSer	255
DB	1275	ACCATCTGCGCTGAGGCTGAAAAATGGGGCGTCAAAACAACAGGTCATCTTGGTATCATG	1334
QY	256	LeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThr	275
DB	1335	ATTGGTGTATGTTTCGGTATCATGTTCTCGAACTATGCTTCGGTTCTCGATGGGATCT	1394
QY	276	SerLeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePhe	295
DB	1395	CGCTTCGTTGTCGGTAAAGAA-----GTCAACGTGGCGCAGTTCGACAGTTTTCATG	1448
QY	296	SerValIleHisSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAla	315
DB	1449	TCATCTCGATCGGTTCTGTTTTCAGTTTGGGCAACGTCGCCCAATGGTCAGGCTTTTACG	1508
QY	316	IleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAsp	335
DB	1509	AATGGTGTCTCGCGCCGCGAAGATTATACACACGATGTGACCCGAGATCGCCACTCGAC	1568
QY	336	AsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsn	355
DB	1569	CCCTATTCTCAGACGGAAGGTACTCGACCATTTTGAAGGAATATCGAATTCGCAAT	1628
QY	356	ValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArg	375
DB	1629	GTCAACACATCTACCTTCAAGACCCGGAAGTTACAGTCATGAGAGATGTTCTCTTTATCG	1688
QY	376	IleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrVal	395
DB	1689	ATGCCGCCGGAAGACTACCGCATTTGGTGGGCCCATCTGGCTCTGGAAGAGTACTGTT	1748
QY	396	ValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsn	415
DB	1749	GTCCGCTTGGTGAGCGCTTTTACTTCCAGTAGGAGGCCAGGATTTCTCTGACGCCCAT	1808
QY	416	AspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGlu	435
DB	1809	GATATCAAAACCTCAACCTTCCTGGCTGGCTGCACAGCAGATCTCTCTGTGACGACGAA	1868

```

Qy 436 ProValLeuPheGlyThrThrIleSerAsnIleLysTyGly----- 450
Db 1869 CCTGCTCTTTTCAGCACCAACGATCTTTAGAAACATCGAACATGGCTTGATTGGCACCAAA 1928
Qy 451 -----ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAla 466
Db 1929 TTCGAGCATGATCGAGGACGAAGATTAGAGAGCTCGTTTGAGAAATGCGGCCAGAAATGGCC 1988
Qy 467 AsnAlaTyAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLys 486
Db 1989 AATGCTCATGATTTTATTATGCTCTGCTGAAGGTTACGATACGAATGTGGTCCAGGT 2048
Qy 487 GlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArg 506
Db 2049 GGTGTTCTTACTTTTCAGGAGGTGAGAAGCAAGCATATVTCGCAATTGCTCGTGCCATTGTGTCAGT 2108
Qy 507 AsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSer 526
Db 2109 GACCCCAAGATTCTGTGCTTATGATGAAGCTACATCAGCTTTTGGATACCAAGTCCGAGGGC 2168
Qy 527 AlaValGlnAlaLeuGluLys 534
Db 2169 GTGTCACAGCGCTCTTGATAAA 2192

RESULT 15
US-08-996-545-1
; Sequence 1, Application US/08996545
; Patent No. 5928898
; GENERAL INFORMATION:
; APPLICANT: Skatrud, Paul L.
; APPLICANT: De Waard, Maarten A.
; APPLICANT: Peery, Robert B.
; APPLICANT: Andrade, Alan C.
; TITLE OF INVENTION: Multiple Drug Resistance Gene atrd of
; TITLE OF INVENTION: Aspergillus Nidulans
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: U.S.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/996,545
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Webster, Thomas D.
; REGISTRATION NUMBER: 39,872
; REFERENCE/DOCKET NUMBER: X-11766
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3334
; TELEFAX: 317-276-2763
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4002 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..4002
US-08-996-545-1

```


Alignment Scores:

Pred. No.:	3,628-97	Length:	4002
Assignment Scores:		Matches:	201
Score:	914.00	Conservative:	103
Percent Similarity:	56.93%	Mismatches:	181
Best Local Similarity:	37.64%	Indels:	50
Query Match:	33.37%	Gaps:	6
DB:	2		

US-09-873-409-7 (1-541) x US-08-996-545-1 (1-4002)

Qy	15	LeuProLeuMetProLeuValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu	34
Db	368	GTGCATTGCTGCGCGTCGACATTTCCAGA-----GGATATGTTAT	409
Qy	35	ValGlnThrAsnThrTyrSerPhePhe-----ArgLeuThrLeuTyrVal	50
Db	410	ATCAAAATCTCGTAGCAC-GAGTTCATGATGAATGACCAAGAACGGTACTGTCTTCGTA	468
Qy	51	GlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeuTrpIleThr	70
Db	469	TACCTCGGTATCGCGAGTTGTGTCACGTGCTATGTTAGTACTGTGGCTTCATCATACC	528
Qy	71	AlaAlaArgGlnThrLysArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAsp	90
Db	529	GGAGAACACCCACGACAGAGATCCGCGAGTATTACCTTGAGTCTTACTCTCGCCGACGAAC	588
Qy	91	IleGlyTrpPheAspSerCysaspIleGlyGluLeuAsnThrArgMetThr--AspIle	109
Db	589	ATTGGCTATTTTGATAACTCGGTGCGCGGGAAGTGACACCCGCTATACAGCCGATACA	648
Qy	110	AspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGlnAsnMetSerThr	129
Db	649	AACCTTATCCAGGATGGCATTTCCGAGAGAGTCGGTCTCACTTTGACTGCCCTGGCGACA	708
Qy	130	PheSerIleGlyLeuAlaValGlyLeuValLysGlyTrpLysLeuThrLeuValThrLeu	149
Db	709	TTCTGTGACAGCATTCATTATCGCTGCTACGTCAAAATACTGGAAGTTGGCTTAAATTG	768
Qy	150	SerThrSerProLeuIleMetAlaSerAlaAlaCysSerArgMetValIleSerLeu	169
Db	769	TCAACAAATTGGCCCTCGTTCACCATGGCGGTGTTCTCAGTTTATCATCAAGTAC	828
Qy	170	ThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluValLeuSer	189
Db	829	AGCAAAAAGTCGCTTGACAGCTACGGTGCAGCGCGCACTGTTGCGAAGAGTCA	888
Qy	190	SerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSerPheLeu	209
Db	889	TCCATCAGAAATGCCACAGCGTTTGGCACCCAAGCAAGCTTGGCAAG-----	936
Qy	210	LeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCysValLeu	229
Db	936	-----936	
Qy	230	***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThr	249
Db	937	-----CAGTATGAGTCCACTTAGACGAAGCTGAGAAATGGGGAAACAAGACACG	987
Qy	250	IleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeu	269
Db	988	ATTGTCATGGGTTTCATGATGGCGCATGTTTGGCCCTTATGTACTCGAACTACGGCTCT	1047
Qy	270	AlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThr	289
Db	1048	GGCTTCTCGAATGGGTTCTCGTTCTTCCTGGTAGATGGT-----GCAGTCGATGTGGGTGAT	1101
Qy	290	ValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValPro	309
Db	1102	ATTCTCACAGTCTCTCATGGCCATCTTGATCGGATCGTCTCTCTGGGAAACGTTAGTCCA	1161
Qy	310	HisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAsp	329
Db	1162	AATGCTCAAGCATTTACAAACCGTGTGGCGGCGCGCAAGATATTTTGGAAACGATCGAT	1221

Search completed: March 31, 2003, 15:37:48
Job time : 148.227 secs

Qy	330	LySLyPProSerIleAspAsnPheserThrAlaGlyTyrLysProGluSerIleGluGly	349
Db	1222	CGCCAGTCCCATTTAGATCCATATTTCGAACGAAGGAAGACGCTCAGCATTTTGAGGCGC	1281
Qy	350	ThrValGluPheIysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeu	369
Db	1282	CACATTGAGTTACGCAATGTCGAAGCATATTTACCACCTTAGACCGGAGGTCAACGTCATG	1341
Qy	370	GLyGLyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGly	389
Db	1342	GAGGATGTTCTCTGTCATGCGCGTGGAAAAACAACCGCTTTAGTCGGCCCTCTGCGC	1401
Qy	390	SerGlyLysSerThrValValGlnLeuGlnArgLeuTyrAspProAspAspGlyPhe	409
Db	1402	TCGGAAGATACGGTGTGCGCTTGGTGAAGCATTTCTACATGCCCTGTGTCGCGGTACG	1461
Qy	410	IleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIle	429
Db	1462	GTTTTGCTGGATGCCCATGACATCAAGNACCTCAATCCCGTGGCTTCGCCAACAGATC	1521
Qy	430	GlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyr	449
Db	1522	TCATTGTTGACGAGGAGCGCTGTTCTTTTGGACACGAGATTTATAAGAAATATTAGGCAC	1581
Qy	450	Gly-----ArgAspAspValThrAspGluGluMetGlu	460
Db	1582	GGTCTCATCGGCACAAGATGACGAGATGAATCCGAGGATAGAGTCCGGGAATCATCTCAG	1641
Qy	461	ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn	480
Db	1642	AACGGGCAAAATGCGAATGCTCATGACTTATTACTGCTTCTGCTTGAAGGTTATGAG	1701
Qy	481	ThrLeuValGlyGlyLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle	500
Db	1702	ACCAATGTTGGGCAGCGTGGCTTCTCTCTTTTCAGGTGGCCAGAAACACGCGCATTTGCAATC	1761
Qy	501	AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu	520
Db	1762	GCCCGTCGGCTGTTAGTACCAACCAAAATCTCTGCTTCCTGGATGAAGCTACTTCGCCCTTG	1821
Qy	521	AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys	534
Db	1822	GACACAAATCCGAAGCGCTGGTTTCAACGAGCTTTGGAGAGG	1863

•

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 31, 2003, 13:52:37 ; Search time 187.092 Seconds
(without alignments)
2337.888 Million cell updates/sec

Title: US-09-873-409-8

Perfect score: 2592
Sequence: 1 MILGILASLVNAGCLFLMPL.....SEKSAVQAALXKTPRYSF 514

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 574371 seqs, 425486471 residues

Total number of hits satisfying chosen parameters: 1148742

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame.p2n.model -DEV=xlh
-Q/cgn2_1/USPTO_spool/US09873409/runat_27032003_115424_19290/app_query.fasta_1.7544
-DB=Published Applications NA -QMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09873409 @CGN 1.1 960 @runat_27032003_115424_19290
-NCPU=6 -ICPU=3 -NO XLPXY -NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	2592	100.0	1940	US-09-873-409-16
2	2568.5	99.1	2021	US-09-873-409-15
3	2552	98.5	3621	US-09-873-409-14
4	2528.5	97.6	3702	US-09-873-409-13

Alignment Scores:			
Pred. No.:	1.14e-267	Length:	1940
Score:	2592.00	Matches:	514
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	10	Gaps:	0

ALIGNMENTS

RESULT 1

US-09-873-409-16
; Sequence 16, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
; APPLICANT: Frank, Markus
; APPLICANT: Sayegh, Mohamed
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
; TITLE OF INVENTION: Homologue on Chromosome 7p15-21 and Uses Thereof
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873,409
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 16
; LENGTH: 1940
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-873-409-16

Sequence 12, Appli
Sequence 5, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 1424, Ap
Sequence 3, Appli
Sequence 1, Appli
Sequence 2, Appli
Sequence 3, Appli
Sequence 1, Appli
Sequence 2, Appli
Sequence 7, Appli
Sequence 2399, Ap
Sequence 1560, Ap
Sequence 483, App
Sequence 11, Appl
Sequence 10, Appl
Sequence 4, Appli
Sequence 3, Appli
Sequence 2337, Ap
Sequence 263, App
Sequence 1, Appli
Sequence 3, Appli
Sequence 2, Appli
Sequence 9, Appli
Sequence 10, Appli
Sequence 1, Appli
Sequence 8, Appli
Sequence 7, Appli
Sequence 9, Appli
Sequence 18, Appli
Sequence 16, Appli
Sequence 15, Appli
Sequence 1500, Ap
Sequence 9908, Ap
Sequence 40, Appl
Sequence 6018, Ap
Sequence 170, App

US-09-873-409-8 (1-514) x US-09-873-409-16 (1-1940)

Qy	1	MetIleLeuGlyIleIeuAlaSerIeuValAsnGlyAlaCysLeuProLeuMetProLeu	20
Db	34	ATGATCTCGGTATACTGGCATCACATGGTCAATGAGAGCTGCCTTCCTTTAATGCCACTG	93
Qy	21	ValLeuGlyGluMetSerAspAenLeuIleSerGlyCysLeuValGlnThrAsnThrTyr	40
Db	94	GTTTTAGAGAAATGAGTGATAACCTTATTAGTGGATGTCTAGTCCAAACTAAACACATAC	153
Qy	41	SerPhePheArgIeuThrIeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGly	60
Db	154	TCCTTCTTCAGGTGACCTGTATTATGTATGGAATAGGTGTTCGTGCCTTCGATTTTGGT	213
Qy	61	TyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLys	80
Db	214	TACATACAGATTTCTCTGTGGATTAACTGCAGCAGCAGACCCAGAGAGATTCGAAAA	273
Qy	81	GlnPhePheHisSerValLeuAlaGlnAspIleIleGlyTrpPheAspSerCysAspIleGly	100
Db	274	CAGTTTTCATTTCAGTTTGGCACAGGACATCGCTGGTTTGATAGCTGTGACATCGGT	333
Qy	101	GluLeuAsnThrArgMetThrAspIleAspLysIleSerAspGlyIleGlyAspLysIle	120
Db	334	GAACCTTAACACTCGCATGACAGACATGACAAAAATCAGTGATGTATTTGGAGATAAGATT	393
Qy	121	AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys	140
Db	394	GCTCTGTGTGTTCAAAACATGCTACTCTTTTCAATGGCTGGCTGGATGGTTGGTGAG	453
Qy	141	GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla	160
Db	454	GGCTGGAAACTCACCTCAGTGTACTCTATCCAGCTCTCTCTTATTAATGGCTTCAGCGCA	513
Qy	161	AlaCysSerArgMetValIleSerIeuThrSerLysGluLeuSerAlaTyrSerLysAla	180
Db	514	GCATGTCTTAGGATGGTCACTCATTTGACCAAGTAGGAATTAAGTGCCCTATTCCAAAGCT	573
Qy	181	GlyAlaValAlaGluValLeuSerSerIleIleArgThrValIleAlaPheArgAlaGln	200
Db	574	GGGCTCTGTCGACAGAAGTCTTGTTCATCAATCCGAACAGTCATAGCTCTTTAGGGCCAC	633
Qy	201	GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyLys	220
Db	634	GAGAAAGAACTTCAAAGGTATACACAGAAATCTCAAAGATGCAAGAGATTTTGGCATAAAA	693
Qy	221	ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr	240
Db	694	AGGACTATAGCTTCAAAGTGCTCTCTTGGTGTGTGTACTTCTTTATGAATGGAACTTAT	753
Qy	241	GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle	260
Db	754	GGACTCTCTTTTGGTATGAACTCTCTTGATTCCTTAATGGAGAACCCTGGATATACCATC	813
Qy	261	GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla	280
Db	814	GGGACTGTCTGCTGTTTCTTTTAGTGTAAATCCATAGACAGTATTGTGATTTGGAGACAGA	873
Qy	281	ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal	300
Db	874	GTCCCTCACTTTGNAACTTCGCAATAGCCCGGAGAGCTGCCTTTCATATTTTCCAGGTT	933
Qy	301	IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle	320
Db	934	ATTGATTAAGAAACCAGTATAGATAACTTTTCCACAGCTGGATATAAAACCTGGAATCCATA	993
Qy	321	GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys	340
Db	994	GAAGGAAGCTGGGAATTTAAAAATGTTTCTTCAATTATCCATCAAGACCATCTATCAAG	1053
Qy	341	IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu	360
Db	1054	ATTCTGAAAGGTCTGAATCTCAGAAATTAAGTCTGGAGACAGCTGCCTGGTGGGTCTC	1113

Qy	361	AsnGlySerGlyLyysSerThrValValGlnLeuLeuGlnAArgLeuTyArgProAsp	380
Db	1114	AATGGCAGTCGGGAAGTAGTACGGTAGTCCAGCTTCTGCGAGAGGTATATGATCCGGATGAT	1173
Qy	381	GlyPheIleMetValAspGluAsnAspIleAArgAlaLeuAsnValAArgHisTyArgAsp	400
Db	1174	GGCTTTATCATGTGGATGAGATGNACATCAGAGCTTTAAATGTGCGCATTTATCGAGC	1233
Qy	401	HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnIle	420
Db	1234	CATATTCCGAGTCGTTAGTCAAGAGCGCTGTTTGTTCGGGACCACCATCAGTAACAATATC	1293
Qy	421	LysTyGlyAArgAspValThrAspGluGluMetGluAArgAlaAArgGluAlaAsn	440
Db	1294	AAGTATCGACGAGATGATGTGACTGATGAAGAGATGGAGAGACGACGAAGGAGCAAAAT	1353
Qy	441	AlaTyArgPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly	460
Db	1354	GCGTATCATTTTATCATGGNGITTCCTAATAAATTTAATACATTCGTAGGGGMAAAGGA	1413
Qy	461	AlaGlnMetSerGlyGlyGlnLysGlnAArgIleAlaIleAlaAArgAlaLeuValAArgAsn	480
Db	1414	GCTCAAAATGAGTCGGAGGCAGAACACAGAGGATCGCAATTTGCTCGTGCCTTAGTTGCAAAAC	1473
Qy	481	ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla	500
Db	1474	CCCNAAGATTCGATTTTATGATNAGGCGTCGTCTGCCCTGGATTCAGAAAGCAGTCAGCT	1533
Qy	501	ValGlnAlaAlaLeuGluLysAspThrProArgTyfyrSerPhe	514
Db	1534	GTTCAAGCTGCATCGGAGGAAGATACCCCCAGGTATTCATTT	1575

RESULT 2

```

US-09-873-409-15
; Sequence 15, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
; APPLICANT: Frank, Markus
; APPLICANT: Savegh, Mohamed
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
; TITLE OF INVENTION: Homologue on Chromosome 7p15-21 and Uses Thereof
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873,409
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
; LENGTH: 7021
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: No. US20020037522A1
; LOCATION: (723)..(723)
; OTHER INFORMATION: n at position 723 represents any nucleotide (A, T, C or G)
US-09-873-409-15

```

Alignment Scores:

Alignment Scores:			
Pred. No.:	3,976-265	Length:	2021
Score:	2568.50	Matches:	514
Percent Similarity:	95.01%	Conservative:	0
Best Local Similarity:	95.01%	Mismatches:	0
Query Match:	99.09%	Indels:	27
DB:	10	Gaps:	1

US-09-873-409-8 (1-514) X US-09-873-409-15 (1-2021)

Qy	1	MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu
Db <td>34</td> <td>ATGATCCTGGGTACTCTGGCATCTGCTCAATGGAGCGTGCCTTCTCTTAATGCCACTG</td>	34	ATGATCCTGGGTACTCTGGCATCTGCTCAATGGAGCGTGCCTTCTCTTAATGCCACTG
Qy <td>21</td> <td>ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr</td>	21	ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr

Db 94 GTTTTAGGAGAAATGAGTGATTAACCTTATTAGTGGGANGTCTAGTCCAAACTAACACATAC 153
Qy 41 SerPhePheArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePheGly 60
Db 154 TCTTTCTTCAGGTTGACCTGTATATATGTTGGAATAGGTGTTGCTGCTTGTATTTTGT 213
Qy 61 TyrIleGlnIleSerLeuTrpIleThrAlaAlaArgGlnThrLysArgIleArgLys 80
Db 214 TACATACAGATTCTCTGTGGATTATAACTGCAGCAGCAGACCAAGAGGATTCCAAAA 273
Qy 81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGly 100
Db 274 CAGTTTTTTTCATTCAGTTTTGGCACAGGACATCGGCTGGTTGATAGCTGTGACATCGT 333
Qy 101 GluLeuAsnThrArgMetThrAspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db 334 GAACCTTAACACTCGCATGACAGACATTGACAAATCAGTGATGTTATCGAGATAAGATT 393
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db 394 GCTCTGTTGTTTCAAAACATGCTACTTTTCGATTGGCTGGCAGTTGGTTGGTGAAG 453
Qy 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db 454 GGCTGGAAACTCACCTAGTGACTCTATCCAGCTCTCTCTTATATATGCTTCAGCGCA 513
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db 514 GCATGTTCTAGGATGGTCACTCATCTGACCAAGTAAGGATTAAGTCCCTATTCCAAAGCT 573
Qy 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db 574 GGCGCTGTGGCAGAGAAGTCTGTGTCATCAATCCGAACAGTCATAGCCTTTAGGGGCCAG 633
Qy 201 GluLysGluLeuGln----- 205
Db 634 GAGAAAGAACTTCAAAGGCTTTCTCTTTAAATATAACAAGATATGCTGGTTTTATTTT 693
Qy 206 -----ArgTyrThrGlnAsnLeuLysAsp 213
Db 694 CCCAGTGGCTACTAAGTGTGTCTGTTNTTGTAAAGGTATACACAGATCTCAAGAT 753
Qy 214 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 233
Db 754 GCAAAGGATTTTGGCATAAAAAGGACTATAGCTTCAAAAGTGTCTCTGCTGTGTGTAC 813
Qy 234 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 253
Db 814 TTCTTTTATGAATGGAACCTATGGACTTGTCTTTTGGTATGGAAACCTCTCTGTATCTTAAT 873
Qy 254 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 273
Db 874 GGAAACCTGGATATACCATCGGAGCTGTCTTCTGCTGTTTCTTTAGTGTATCCATAGC 933
Qy 274 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 293
Db 934 AGTTATTGCAATGGAGCAGCAGTCCCTCACTTTGAAACCTTCGCAATAGCCCGAGAGCT 993
Qy 294 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 313
Db 994 GCCTTTTCATATTTTCCAGGTTATTGATAAGAAACCCAGTATAGATAACTTTTCCAGCT 1053
Qy 314 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 333
Db 1054 GGATATAAACCTGAATCCATAGAGGAACGTGTGGAATTTTAAAAATGTTTCTTTCAATTAT 1113
Qy 334 ProSerArgProSerIleLysIleLeuLysGlyLysAsnLeuArgIleLysSerGlyGlu 353
Db 1114 CCATCAAGACCATCTATCAAGATTCTGAAGGCTCTGAATCTCAGATTAAAGTCTGGAGAG 1173
Qy 354 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 373
Db 1174 ACAGTGCCTTGGTCTCTCAATGGCAGTGGGAAGAGTAGTCGCTAGTCCAGCTTCTGCAG 1233

Qy 374 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 393
Db 1234 AGGTTATATGATCCGGATGATGGCTTTTATCATGGTGGATGAGATCAGATCAGAGCTTTA 1293
Qy 394 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 413
Db 1294 AATGTGCGCATTTATCGAGACCATATTGGAGTGGTTAGTCAAGAGCCTGTTTGTTCGGG 1353
Qy 414 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 433
Db 1354 ACCACCATCAGTAACATATCAAGTATGGACGAGATGATGTGCTGATGATGAGATGGAG 1413
Qy 434 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 453
Db 1414 AGAGCAGCAAGGGAAGCAAAATCGCTATGATTTTATCATGGAGTTCCTTAATAAATTAAT 1473
Qy 454 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 473
Db 1474 ACATTTGGTAGGGGNAAGGAGCTCAAAATGAGTGGAGGGCAGAAACAGAGGATCGCAATT 1533
Qy 474 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 493
Db 1534 GCTCGTGCCTTAGTTTGAACCCCAAGATTCTGATTTTAGATGAGGCTACGTTGCCCTG 1593
Qy 494 AspSerGluSerLysSerAlaValGlnAlaLeuGluLysAspThrProArgTyrSer 513
Db 1594 GATTTCAGAAAGCAAGTCAGCTGTTCAAGCTGCACTCGAGAGGATACCCCCAGGTATTCA 1653
Qy 514 Phe 514
Db 1654 TTT 1656
RESULT 3
US-09-873-409-14
; Sequence 14, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
; APPLICANT: Frank, Markus
; APPLICANT: Sayegh, Mohamed
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873,409
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
; LENGTH: 3621
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-873-409-14
Alignment Scores:
Pred. No.: 5,61e-263 Length: 3621
Score: 2552.00 Matches: 507
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.46% Indels: 0
DB: 10 Gaps: 0
US-09-873-409-8 (1-514) x US-09-873-409-14 (1-3621)
Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 34 ATGATCCTGGGTATATCTGGCATCCTGGTCAATGGAGCCTGCCCTCTTTTAATGCCACTG 93
Qy 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAnthrTyr 40
Db 94 GTTTTAGGAGAAATGAGTGTAACTTATTAGTGGATGTCTAGTCCAACTAACACATAC 153
Qy 41 SerPhePheArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePheGly 60


```
Db 274 CAGTTTTTTCATTCACTGTTGGCACAGGACATCGGCTGGTGTGATAGCTGTGCATCGGT 333
Qy 101 GluLeuAsnThrArgMetThrAspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db 334 GAACTTTAACACATCGCATGCAGACAGCATTTGACAAATCAGTGTGATTTGGAGATAAGATT 393
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db 394 GCTCTGTTGTTTCAAAACATGCTACTTTTTCGATTGGCTGCGCATTGGTTGGTGAAG 453
Qy 141 GlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db 454 GCGTGAATCTCACCTAGTACTCTATCCAGCTCTCTCTTAATAGGCTTCAGCGCA 513
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db 514 GCATGTTCTAGGATGCTCATCTCATTTGACCAGTAAGGAATTAAGTGCCTATTCCAAAGCT 573
Qy 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db 574 GGGGCTGTGGCAGAAAGTCTTGTCATCAATCCGAACAGTCATAGCCTTTAGGGGCCAG 633
Qy 201 GluLysGluLeuGln----- 205
Db 634 GAGAAAGAACTTCAAAGGCTTTTCCTTTTAAATATAACAAGATATGCTGGTGTATTTT 693
Qy 206 -----ArgTyrThrGlnAsnLeuLysAsp 213
Db 694 CCCAGTGGCTACTAAGTGTGTCTGTTNTTGTAAAGTATACAGAAATCTCAAGAT 753
Qy 214 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 233
Db 754 GCAAAAGATTTTGGCATAAAAGGACTATAGCTTCAAAAGTGTCTCTGTGTGTGTAC 813
Qy 234 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsn 253
Db 814 TTTCTTTATGAATGGAACTTATGGACTTGCTTTTGTATGGAACCTCTTGATTTTAAT 873
Qy 254 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 273
Db 874 GGGAACCTGGATATACCATCGGACTGTCTTGTGTTTCTTTAGTGTATCCATAGC 933
Qy 274 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 293
Db 934 AGTTATTGCTATGGACAGCAGTCCCTCACTTTGAAACCTTCGCAATAGCCCGAGGAGCT 993
Qy 294 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 313
Db 994 GCCTTTTCATATTTTCCAGGTTATTGTATAAGAAACCCAGTATAGATAACTTTTCCAGCT 1053
Qy 314 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 333
Db 1054 GGATATAAACCTGAATCCATAGAAGGAACGTGTGGAATTTTAAATGTTTCTTTCAATTAT 1113
Qy 334 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 353
Db 1114 CCATCAAGACCATCTATCAAGATCTTGAAAGTCTGAATCTCAGAAATTAAGTCTGGAG 1173
Qy 354 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 373
Db 1174 ACAGTGGCTTGGTGGTCTCAATGGCAGTGGGAAGATACGGTAGTCTCCAGCTTCTGCAG 1233
Qy 374 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 393
Db 1234 AGGTTATATGATCCGATGATGGCTTTATCATGTTGGATGAGAAATGACATCAGAGCTTTA 1293
Qy 394 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 413
Db 1294 AATGTGGGCAATTATCGAGACCATATTTGAGTGGTGTAGTCAAGAGGCTGTTTGTTCGGG 1353
Qy 414 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluLeuMetGlu 433
Db 1354 ACCACCATCAGTAACAATATCAAGTATGGACGAGATGATGTGACTGATGAAGATGGAG 1413
```

```
Qy 434 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 453
Db 1414 AGAGCAGCAAGGGAAGCAAAATCGGTATGATTTTATCATGGAGTTTCTTAATAAATTAAT 1473
Qy 454 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 473
Db 1474 ACATTGGTAGGGGAAAAGGAGCTCAAAATGATGGAGGCGAAGACAGAGGATCGCAATT 1533
Qy 474 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 493
Db 1534 GCTCGTGCTTAGTTTGGAAACCCCAAGATTTCTGATTTTAGATGAGGCTACGCTGCCCTG 1593
Qy 494 AspSerGluSerLysSerAlaValGlnAlaLeuGluLys 507
Db 1594 GATTTCAGAAAGCAAGTCAGCTGTTCAAGCTGCACCTGGAGAAG 1635
```

RESULT 5

```
US-09-873-409-12
; Sequence 12, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Frank, Markus
; APPLICANT: Sayegh, Mohamed
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
; FILE OF INVENTION: Homologue on Chromosome 7p15-21 and Uses Thereof
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873,409
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 3177
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: No. US20020037522A1e
; LOCATION: (198)..(198)
; OTHER INFORMATION: n at position 198 represents any nucleotide (A, T, C or G)
US-09-873-409-12
```

```
Alignment Scores:
Pred. No.: 8,09e-172 Length: 3177
Score: 1699.50 Matches: 343
Percent Similarity: 92.70% Conservative: 0
Best Local Similarity: 92.70% Mismatches: 0
Query Match: 65.57% Indels: 27
DB: 10 Gaps: 1
```

US-09-873-409-8 (1-514) x US-09-873-409-12 (1-3177)

```
Qy 165 MetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAla 184
Db 1 ATGGTCAATCTCATTTGACCAAGTAAAGAAATTAAGTGCCTATTCCAAAGCTGGGCTGTGGCA 60
Qy 185 GluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeu 204
Db 61 GAAGAAGTCTTGTCAATCCGACAGTCATAGCCTTTAGGGCCCCAGGAGAAGACTT 120
Qy 205 Gln----- 205
Db 121 CAAAGGCTTCTCTTTTAAATATAACAAGATATGCTTGGTTTATTTTCCCCAGTGGCTA 180
Qy 206 -----ArgTyrThrGlnAsnLeuLysAspAlaLysAspPhe 217
Db 181 CTAAGTGTGTCTGTGTTNTTGTAGGTATACACAAATCTCAAGATGCAAGAGATTTT 240
Qy 218 GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsn 237
Db 241 GGCATTAAGAGCACTATAGCTTCAAAAGTGTCTCTTGGTGTGTGTACTTCTTTATGAAT 300
Qy 238 GlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsnGlyGluProGly 257
```


Db 301 GGAACCTATGGACTGCTTTTGGTATGGAACCTCTTGTATCTTAATGAGAACCTGGA 360
Qy 258 TyThrIleGlyThrValLeuAlaValPheSerValIleHisSerSerTyrCysIle 277
Db 361 TATACCATCGGAGCTGTTCTTGTGCTTTCTTAGTGTATTCATAGCAGTATTGCAAT 420
Qy 278 GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 297
Db 421 GGAGCAGCAGTCCCTCACTTTGAAACCTTCGCAATAGCCCGAGAGCTGCTTTCATATT 480
Qy 298 PheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysPro 317
Db 481 TTCAGGCTTATGATAAGAAACCCAGTATAGATAAATCTTCCACAGCTGGATATAAACCT 540
Qy 318 GluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgPro 337
Db 541 GAATCCATAGAGGAACCTGTGGAATTTAAATAATGTTCTTCAATATCATCATCAAGACCA 600
Qy 338 SerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu 357
Db 601 TCTATCAAGATTCTGAAGGCTCTGAATCTCAGAAATTAAGTCTGGAGAGCAGTCGCTTG 660
Qy 358 ValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAsp 377
Db 661 GTCGGTCTCAATGCGCAGTGGGAGAGTACGCTAGTCCAGCTTCTGCAGAGGTTATATGAT 720
Qy 378 ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis 397
Db 721 CCGGATGATGGCTTTATCATGGTGGATGAGAAATGACATCAGAGCTTTAAATGTGCGGCAT 780
Qy 398 TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSer 417
Db 781 TATCGAGACCATATTGGAGTGGTTAGTCAAGAGCTGTTTGTTCGGGACCAACCATCAGT 840
Qy 418 AsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArg 437
Db 841 AACAAATATCAATGATGACGAGATGATGCTGATCAAGAGATGGAGAGCAGCAAGG 900
Qy 438 GluAlaAsnAlaTyrAspPheIleMetGluPheProLeuLysPheAsnThrLeuValGly 457
Db 901 GAGCAAAATGCGTATGATTTTATCATGGAGTTTCCATAAATTAATATCATTTGGTAGGG 960
Qy 458 GluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu 477
Db 961 GAAAAAGAGCTCAAATGAGTGGAGGCGAGAACAGAGGATCGCAATGCTCGTGCCTTA 1020
Qy 478 ValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 497
Db 1021 GTTCGAAACCCCAAGATCTGATTTTATGATGAGGCTACGCTGCGCTGGATTCAGAAAGC 1080
Qy 498 LysSerAlaValGlnAlaAlaLeuGluLys 507
Db 1081 AAGTCAGCTGTTCAAGCTGCACTGGAGAG 1110

RESULT 6
US-09-866-866A-5
; Sequence 5, Application US/09866866A
; Patent No. US20020102244A1
; GENERAL INFORMATION:
; APPLICANT: Sorrentino, Brian
; APPLICANT: Schuetz, John
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
; FILE REFERENCE: 1340-1-021CIP2
; CURRENT APPLICATION NUMBER: US/09/866, 866A
; PRIOR FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 09/584,586
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: PCT/US99/11825
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: 60/086,988
; PRIOR FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 5
; LENGTH: 4189
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-866-866A-5

Alignment Scores:
Pred. No.: 6,47e-160 Length: 4189
Score: 1590.00 Matches: 307
Percent Similarity: 76.32% Conservatives: 99
Best Local Similarity: 57.71% Mismatches: 94
Query Match: 61.34% Indels: 32
DB: 10 Gaps: 5

US-09-873-409-8 (1-514) x US-09-866-866A-5 (1-4189)
Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 148 ATGATTCTGGGAACCTCTCGCTGCTATTATTCATGGAACATTACTTCCCTCTTGTGCTG 207
Qy 21 ValLeuGlyGluMetSerAspAsnLeuIle----- 30
Db 208 GTCTTTGGAAACATGACAGATAGTTTACAAAGCAGAGCCAGTATTCTGCCAAGCAATT 267
Qy 31 -----SerGlyCysLeuValGlnThrAsnThr----- 39
Db 268 ACTAATCAAAAGTGGACCCCAACAGTACTCTGATCATCAGCAACAGCAGTCTGGAGGAGAG 327
Qy 40 -----TyrSerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAla 56
Db 328 ATGCCATATACGCCCTAC-----TATTACACCGGATTTGGTGTGTGTG 372
Qy 57 LeuIlePheGlyTyrIleGlnIleSerLeuTyrIleThrAlaAlaArgGlnThrLys 76
Db 373 CTCTATGTTGCCCTACATCCAGGTTTCACTTTGGTGGCTGGCAGCTGGAAGACAGATACAC 432
Qy 77 ArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSer 96
Db 433 AAGATTAGGCAAGATTTTCCATGCTATAATGAATCAGGAGATAGCTGGTGTGTG 492
Qy 97 CysAspIleGlyLeuLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGly 115
Db 493 CATGATGTTGGGAGCTCAACACCCGCTCACAGATGATGCTCTCCAAATTAATAGCGGA 552
Qy 116 IleGlyAspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAla 135
Db 553 ATGGTGACAAAATGGGATGTTTTCAGTCCATCAACACCATTTTACCGCTTTAGCCGCTTTATC 612
Qy 136 ValGlyLeuValLysGlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIle 155
Db 613 ATAGGATTTAAGTGGTTGGAGCTAACCTTTGTCATTTTGGCTGTCAGCCCTCTTATT 672
Qy 156 MetAlaSerAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSer 175
Db 673 GGATTGTCATCTGCTTTGGGCAAGGATTTGACTTCATTTACTAATAAGGAACCTCCAG 732
Qy 176 AlaTyrSerLysAlaGlyAlaValAlaGluValLeuSerSerIleArgThrValIle 195
Db 733 GCTTATGCAAAAGCTGGAGCGATTTGCTGAAGAGTCTTAGACCCATCAGAACTGTGATT 792
Qy 196 AlaPheArgAlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLys 215
Db 793 GCCTTTGGAGGACACAGAGGAGACTTGAAGGTACATAAATAATTTAGAGAGCTAAA 852
Qy 216 AspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePhe 235
Db 853 AATGTTGGCATAAAGAAAGCTATCACAGCCAGCATTTTCGATAGGCATTTGCTACCTGTTG 912
Qy 236 MetAsnGlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsnGlyGlu 255
Db 913 GTCTATGCATCATATGCATGGCAATTTCTGTATGGGACATCTCTTGTGCTCTCAAAATGAA 972
Qy 256 ProGlyTyrThrIleGlyThrValLeuAlaValPheSerValIleHisSerSerTyr 275

QY 281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
Db 1045 TCTCCAAGCATTTGAAGCATTTGCAATGCAAGAGGAGCAGCTTATGAATCTTCAAGATA 1104
QY 301 IleAspIysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
Db 1105 ATTGATTAATGAAGCAAGTATTGACAGTATTGCAAGAGTGGGCACAAACAGATAATATT 1164
QY 321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
Db 1165 AAGGAATTTGGNAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAGAGTTAAG 1224
QY 341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
Db 1225 ATCTTGAAGGCTGTAACCTGAAGGTGCAGAGTGGCAGCGGTGGCCCTGGTTGGAAC 1284
QY 361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
Db 1285 AGTGGCTGTGGGAAGACACACAGTCCAGCTGATGCAGAGGCTCTATGACCCACAGAG 1344
QY 381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
Db 1345 GGGATGTGAGTGTGATGAGCAGAGATATTAGGACCAATATGTAAGTTCCTACGGGA 1404
QY 401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
Db 1405 ATCATTGGTGTGAGTGCAGGAACCTGTATTGTTGCCACCCAGATAGTGAACATTT 1464
QY 421 LysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
Db 1465 CGCTATGGCGTGAAATGTCCACATGGATGAGATTGAGAAGCTGTCAGAGGAAGCCAA 1524
QY 441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGly 460
Db 1525 GCCATGACCTTATCATGAACCTCCCTCATAAATTTGACACCCCTGGTTGAGAGAGGG 1584
QY 461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db 1585 GCCAGTTGAGTGTGGCAGAACACAGAGATGCCATTCAGCTGCCCTGGTTGCGCAAC 1644
QY 481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db 1645 CCCAAGATCTCTCTGCTGGATGAGGCCACGTCAGCTTGGACACAGAAAGCGAAGCAGT 1704
QY 501 ValGlnAlaAlaLeuGluLys 507
Db 1705 GTTCAGGTGGCTCTGGATAAG 1725

RESULT 8

US-10-044-671-1
; Sequence 1, Application US/10044671
; Patent No. US20020177147A1

GENERAL INFORMATION:

; APPLICANT: Washington State University Research Foundation
; APPLICANT: Mealey, Katrina
; APPLICANT: Bentjen, Steven
; TITLE OF INVENTION: MDRI VARIANTS AND METHODS FOR THEIR USE
; FILE REFERENCE: 4630-61733
; CURRENT APPLICATION NUMBER: US/10/044,671
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/261,578
; PRIOR FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: US 60/314,829
; PRIOR FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1

SEQ ID NO 1

LENGTH: 4317

TYPE: DNA

ORGANISM: Canis familiaris

FEATURE:

NAME/KEY: CDS

LOCATION: (70)..(3912)

OTHER INFORMATION:

US-10-044-671-1

Alignment Scores:

Pred. No.: 2,13e-158 Length: 4317
Score: 1576.00 Matches: 302
Percent Similarity: 75.61% Conservative: 98
Best Local Similarity: 57.09% Mismatches: 105
Query Match: 60.80% Indels: 24
DB: 9 Gaps: 4

US-09-873-409-8 (1-514) x US-10-044-671-1 (1-4317)

QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyValaCysLeuProLeuMetProLeu 20
Db 217 ATGTTGGTGGGACATGGCTGCCATCATCCATGGAGTGCACCTCCTCTCATGATGCTG 276
QY 21 ValLeuGlyCyluMetSerAspAsnLeuIleSerGlyCysLeu----- 34
Db 277 GTTTTGGAAACATGACAGATAGCTTTGCAATGTCAGGAATTTTCAAGAAACAAACTTTT 336
QY 35 -----ValGlnThrAsnThrTyrSerPhePheArg----- 44
Db 337 CCAGTTATAATTAATGAAGTATTACGAACAATACACATTCATCAACCATCTGGAG 396
QY 45 -----LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe 59
Db 397 GAGGAATGACCACGATATGCCATATTATACAGTGGGATCGGTGCTGGCGTGGTGGCT 456
QY 60 GlyTyrIleGlnIleSerLeuTyrIleIleThrAlaAlaArgGlnThrLysArgIleArg 79
Db 457 GCTTACATCCAGGTTCATTCTGGTGGTGGCAGCAGGAGACAGATACTCAAAATATGA 516
QY 80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIle 99
Db 517 AAACAATTTTTCATGCTATCATGCGACAGGAGATTGGCTGGTTTACGTGCATAGCATT 576
QY 100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
Db 577 GGGGAGCTTAACACCGGCTCACAGACGATGCTCCAAATCAATGAAGGAATGGGAC 636
QY 119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
Db 637 AAAGTTGGAATGTCTTTCAATCAATAGCAACATTTTTCACCGGTTTATATAGTGGGTTT 696
QY 139 ValLysGlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
Db 697 ACACCTGGTTGGAAGCTTAACCTTGTGATTGTGGCCATCAGCCCTGTTCTTGACATTTC 756
QY 159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
Db 757 GCGGCCATCTGGGCAAGATAGTACTATCTTCAATTTACTGATAAAGAACTCTTGGGCTATGCA 816
QY 179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198
Db 817 AAAGCTGGAGCAGTAGCTGAAGAAAGTCTTAGCAGCAATCAGAACTGTGATGCTTTGGA 876
QY 199 AlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 218
Db 877 GGCACAAAGAAAGAACTTGAAGGTACAAACAAATTTTAGAAGAGCTAAAGAAATTTGGG 936
QY 219 IleLysArgThrIleAlaSerLysValSerLeuGlyValaValTyrPhePheMetAsnGly 238
Db 937 ATAAAGAAAGCTATACGGCCACATTTCTATTGGTGGCGCTTCTTATTGATCATGCA 996
QY 239 ThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 258
Db 997 TCATATGCTCTGGCTTCTGGTATGGGACCTCCTTGGTCTCTCCAGTGA-----TAT 1050
QY 259 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 278
Db 1051 ACTATTGGACAGGTACTACTGCTCTTCTTTCTGTATTAAATTGGGGCTTTTAGTATTGGA 1110

QY 279 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 298
Db 1111 CAGGCATCCCAAGCATTTGAAGCATTTGCAACGCGAGAGGAGCAGCTTATGAATCTTC 1170
QY 299 GlnValIleAspLysPheProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 318
Db 1171 AGATATTTGACAAATTAACCAACGATTCAGAGCTATTCGAGAGTGACATAACAGAT 1230
QY 319 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 338
Db 1231 AATATTAAAGGAATTTGGAATTCAAATATGTTCACTTCAGTTACCTTCTCGAAGAA 1290
QY 339 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 358
Db 1291 GTTAAGATCTTAAGAGGTCTCAACCTGAAGGTTTCAGAGTGGCGAGCAGTGGCGTGT 1350
QY 359 GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro 378
Db 1351 GGAACAGTGGCTGCGGGAAGAGCAGCGCTGAGCTGATGAGAGGCTCTATGACCCC 1410
QY 379 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 398
Db 1411 ACAGATGGCATGCTGCTATTGATGACAGGACATTAGGACCAATAATGTAAGGCATCTT 1470
QY 399 ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn 418
Db 1471 CGGGAATTAAGTCTGCTGAGTCAAGGAGCTGTGTTGTTGCCACCAAGATAGCTGAA 1530
QY 419 AsnIleLysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGlu 438
Db 1531 AACATTGCTATGCGCGCGAATAATGTCATCGATGAGATGAGATGAGAAAGCTGTTAAGGA 1590
QY 439 AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 458
Db 1591 GCCAATGCTATGATTTATCATGAACACTACCTAATAATTTGACACTCTGGTGGAGAG 1650
QY 459 LysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaLeuVal 478
Db 1651 AGAGGGGCGCGGTGAGTGGTGACAGAAACAGAGAATCGCCATTGTCGGGCGCTGGTT 1710
QY 479 ArgAsnProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 498
Db 1711 CGCAACCCCAAGATTTCTGCTGGAATGAGGCAACGTCAGCTCGACACTGGAAGTGA 1770
QY 499 SerAlaValGlnAlaLeuGluLys 507
Db 1771 GCAGTGGTTCAGTGGCTTCGATTAAG 1797

RESULT 9

US-09-917-800A-1424
; Sequence 1424, Application US/09917800A
; Patent No. US20020119462A1
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Castle, Arthur
; APPLICANT: Elashoff, Michael
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5038-US
; CURRENT APPLICATION NUMBER: US/09/917,800A
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/222,040
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: US 60/222,880
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: US 60/290,029
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/290,645
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: US 60/292,336
; PRIOR FILING DATE: 2001-05-22

; PRIOR APPLICATION NUMBER: US 60/295,798
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/297,457
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,884
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,459
; NUMBER OF SEQ ID NOS: 1740
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1424
; LENGTH: 4254
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020119462A1 M81855
US-09-917-800A-1424
Alignment Scores:
Pred. No.: 6,32e-158 Length: 4254
Score: 1571.50 Matches: 307
Percent Similarity: 75.95% Conservative: 94
Best Local Similarity: 58.14% Mismatches: 104
Query Match: 60.63% Indels: 23
DB: 10 Gaps: 4
US-09-873-409-8 (1-514) x US-09-917-800A-1424 (1-4254)
QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 247 ATGGCTCTGGGAACCTCTCGCTGCTATCATCCACGGAACCCCTGCTTCCCTCTCTGATGCTG 306
QY 21 ValLeuGlyGluMetSerAspAsnLeuIle----- 30
Db 307 GTGTCGGATACATGACATAGATTACCCCAAGCAGAGACCCGCACTTCGACCGAGCG 366
QY 31 -----SerGlyCysLeuValGlnThrAsnThrTyrSerPhePheArgLeu----- 45
Db 367 ATTACTAATCAAAAGTGAATCAACAGTACATACATACCGTCAGCGACACAGATCTCGAGGAG 426
QY 46 -----ThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGly 60
Db 427 GACATGGCCATGTATGCCCTACTATTACACGGGCAATGGTGGCGGTGTGCTCATCGTTGCC 486
QY 61 TyrIleGlnIleSerLeuTyrIleIleThrAlaAlaArgGlnThrLysArgIleArgLys 80
Db 487 TACATCCAGTTTCACTTTGGTGGCTGGCAGTGGGAGACAAATACACAGATTAGGCAG 546
QY 81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGly 100
Db 547 AAGTTTTTCCATGCCATCATGAATCAGGAGATAGGCTGTTTGACGTGAATGACGCTGGG 606
QY 101 GluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLys 119
Db 607 GAGCTCAACACCGCTCTCACAGATGACGTCTCCAAATTAATAGCGGAATTTGGTGCAAA 666
QY 120 IleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuVal 139
Db 667 CTTGGAATGTTCTTTCAGTCCATACAGCATTTTACCGCGTTTATATAGATTTATA 726
QY 140 LysGlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAla 159
Db 727 AGTGGTTGGAAGCTAACCTTGTAAATTTTGGCGCTCAGCCCTCTTATTTGGTTGTCTCT 786
QY 160 AlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLys 179
Db 787 GCCATGTGGGCAAGGTACTGCTTCAATTACTTAATPAAGGAACCTCCAGGCTTATGCAAA 846
QY 180 AlaGlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAla 199
Db 847 GCTGGAGCAGTTGCCCAAGAAAGTCTTAGCAGCCATCAGAACTGTGATTTGGTTTGAGGA 906
QY 200 GlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIle 219


```

Db 907 CAAAAGAGAACTTGAAGGTAACAATAAAATTTAGAAAGCTAAAGAGTTGGCATA 966
Qy 220 LysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPheMetAsnGlyThr 239
Db 967 AAGAAAGCCATCAGCGCAACATTTCCATAGGTATTGCGCTACCTGTGGTCTATGCGTCT 1026
Qy 240 TyrGlyLeuAlaPheTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThr 259
Db 1027 TATGCACTGGCATTTCTGGTATGGACCTCTTGGTCTCTCAATGAA-----TATTCT 1080
Qy 260 IleGlyThrValLeuAlaValPhePheSerValIleHisSerTyrCysIleGlyAla 279
Db 1081 ATTGGCAAGTGTCTACCGTCTCTCTCTATTTATTATGGGACTTTTCAGTATTGGACAT 1140
Qy 280 AlaValProHisPheGluThrPheAlaIleAlaAArgGlyAlaIlePheHisIlePheGln 299
Db 1141 TTAGCCCAACATAGAAAGCTTTTGCAAATGCAAGAGGGCGACCTATGAAATCTTCAAG 1200
Qy 300 ValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSer 319
Db 1201 ATAATTGATTAATGAGCCACAGCATCGACAGTCTTCAACCAAGGACACACAAACAGACAGT 1260
Qy 320 IleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIle 339
Db 1261 ATAATGGCAAAATTTGGAATTTAAAAATGTTTACTTCAACTACCCATCACCAGAGTGAAGTT 1320
Qy 340 LysIleLeuLysGlyLeuAsnLeuAArgIleLysSerGlyGluThrValAlaLeuValGly 359
Db 1321 AAGATCTTGAAGGGCCCTCAACCTGAAGCGTGAAGAGCGGGCAGACGGTAGCCCTGGTGGC 1380
Qy 360 LeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAsp 379
Db 1381 AACAGTGGCTGTGGAAAGACACAACTGTCAGCTGCTGCAGAGGCTCTACGACCCCATTA 1440
Qy 380 AspGlyPheIleMetValAspGluLysAsnAspIleArgAlaLeuAsnValArgHisTyrArg 399
Db 1441 GAGGGCAGGTCAGTATTGACGGACAGGACATCAGGACCATCAATGTGAGGTATCTCGCG 1500
Qy 400 AspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsn 419
Db 1501 GAAATCATTTGGGGTGGTGAAGTCAGGAACCCGCTGCTGTTGCCACCAAGATGCGGAAAC 1560
Qy 420 IleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaAArgGluAla 439
Db 1561 ATTCGCTATGCCGAGAAACCTCACCATGATGATAGAGAAGCTGTCAAGGAAGCC 1620
Qy 440 AsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLys 459
Db 1621 AATGCCTATGACTTCATCATGAAACTGCCCCACAAATTTGACACCCCTGGTTGGTGAGAGA 1680
Qy 460 GlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaAArgAlaLeuValArg 479
Db 1681 GGGCGCAGCTAGTGGGGGACAGAAACAGAGGATCGCCATGCCCCGGGCCCTGGTCCGC 1740
Qy 480 AsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSer 499
Db 1741 AACCCCAAGATCTCTTTGTTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCC 1800
Qy 500 AlaValGlnAlaAlaLeuGluLys 507
Db 1801 GTGGTTCAGGCGCTCTGGATAAG 1824
```

RESULT 10

```

US-09-866-866A-3
; Sequence 3, Application US/09866866A
; Patent No. US2002010224A1
; GENERAL INFORMATION:
; APPLICANT: Sorrentino, Brian
; APPLICANT: Schuetz, John
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
; FILE REFERENCE: 1340-1-021CIP2
; CURRENT APPLICATION NUMBER: US/09/866,866A
```

```

; CURRENT FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 09/584,586
; FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: PCT/US99/11825
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: 60/086,988
; PRIOR FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 3860
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-866-866A-3

Alignment Scores:
Score: 7,89e-158 Length: 3860
Percent Similarity: 1570.00 Matches: 305
Best Local Similarity: 76.85% Conservatives: 100
Query Match: 57.87% Mismatches: 100
DB: 10 Indels: 22 Gaps: 5
```

US-09-873-409-8 (1-514) x US-09-866-866A-3 (1-3860)

```

Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 151 ATGGTGGTGGGAACCTTGGCTGCCATCATCCATGGGGCTGGAGCTTCTCTCATGATGCTG 210
Qy 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
Db 211 GTGTTTGGAGAAATGACAGATATCTTTGCAAAATGCAGGAAATTTAGAGAGATCTGATGTC 270
Qy 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
Db 271 AACATCACTAATAGAAAGTGATATCAATGATGATACAGGGTCTTTCATGATCTGGAGGAGAC 330
Qy 44 -----ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db 331 ATGACCAGATATCCCTATTATTACAGTGGAAATGGTGGGTGCTGGTGTGCTGTAC 390
Qy 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaAArgGlnThrLysArgLysGln 81
Db 391 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 450
Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101
Db 451 TTTTTCATGCTATAATGCGACAGGAGATAGGCTGTTTGTGTCGACCATGTGGGGAG 510
Qy 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db 511 CTTAACACCCGACTTACAGATGATGCTCTTAAGATTAAATGAAGTTATTGGTGACAAAATT 570
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db 571 GGAAGTGTCTTTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGGATTTACACCT 630
Qy 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db 631 GGTGGAAGCTAAACCTTGGATTGTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 690
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db 691 GTCTGGCAAGATACACTACTTCTTCTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 750
Qy 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db 751 GGAGCAGTAGCTCAAGAGGCTCTTGGCAGCAATTTAGAACTGTGATTGCTTTCATTTGGAGCAA 810
Qy 201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
Db 811 AAGAAAGAACTTGAAGGTTACAAACAAATTTTAGAAGAGCTAAAGAAATTTGGGATTAAG 870
```



```
QY 221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAenGlyThrTyr 240
DB 871 AAAGCTATTACAGCAATATTCTATAGTGTCTCTCTGTGATCTATGTCATCTTAT 930
QY 241 GlyLeuAlaPheTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
DB 931 GCTCTGGCTTCTGGTATGGGACCCTTGGTCTCTCAGGGGAA-----TATTCTATT 984
QY 261 GlyThrValLeuAlaValPhePheSerValIleHisSerTyrCysIleGlyAlaAla 280
DB 985 GCACAGTACTCACTGTATTCTTCTGTATTAAATGGGGCTTTAGTGTGGCAGGCA 1044
QY 281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
DB 1045 TCTCCAAGCATTGAAGCAATTTGCAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1104
QY 301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
DB 1105 ATTGATAATAAGCAAGTATTGACAGCTATTGCAAGAGTGGGCACAAACAGATAATATT 1164
QY 321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
DB 1165 AAGGGAATTTGGAATTCAGAAATGTTCACTTCAGTACCCTATCCATCTCGAAAGATTAG 1224
QY 341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
DB 1225 ATCTTGAAGGGCTGAACCTGAAGGTGCAGAGTGGGCAGAGGTGGCCCTGGTTGGAAC 1284
QY 361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
DB 1285 AGTGGCTGTGGGAAGAGCAACAGTCCAGCTGATGCAGAGGCTCTATGACCCACAGAG 1344
QY 381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
DB 1345 GGATGTGTCAGTGTGATGCAGAGGATATTAGACCATTAATGTAAGTTTCTACGGGAA 1404
QY 401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnIle 420
DB 1405 ATCATTTGGTGTGTGATGCAGGAACCTGTATTGTTGCCACCACGATAGTGAACATTT 1464
QY 421 LysTyrGlyArgAspValThrAspGluLeuGluMetGluArgAlaAlaArgGluAlaAsn 440
DB 1465 CGCTATGGCCGTGAAATGTCAACCATGTGATGAGATGAGATGAGAAAGCTGTCAAGGAAGCCAAT 1524
QY 441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
DB 1525 GCTATGACTTTATCAATGAAACTGCTCTCAATAATTTGACACCTGTTGGAGAGAGGG 1584
QY 461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
DB 1585 GCCAGTTGAGTGGTGGCAGAGCAGAGGATGCCATTGCACGTGCCCTGGTTGCAAC 1644
QY 481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
DB 1645 CCCAAGATCCTCTGCTGATGAGGCCACGTCAGCTTGGACACAGAAAGCGAAGCAGTG 1704
QY 501 ValGlnAlaAlaLeuGluLys 507
DB 1705 GTTCAGGTGGCTCTGGATAAG 1725
```

RESULT 11

```
US-09-306-417-1
; Sequence 1, Application US/09306417
; Patent No. US20020103144A1
; GENERAL INFORMATION:
; APPLICANT: Heinrich-Pette-Institut
; TITLE OF INVENTION: Retroviral Gene Transfer Vectors
; FILE REFERENCE: P50491
; CURRENT APPLICATION NUMBER: US/09/306,417
; CURRENT FILING DATE: 1999-05-06
; EARLIER APPLICATION NUMBER: DE 198 22 115
; EARLIER FILING DATE: 1998-05-08
; NUMBER OF SEQ ID NOS: 17
```

```
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 8630
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: proviral
; OTHER INFORMATION: plasmid DNA
```

```
; NAME/KEY: misc feature
; LOCATION: (1)-(160)
; OTHER INFORMATION: plasmid backbone (pUC)
```

```
; NAME/KEY: misc feature
; LOCATION: (161)..(677)
; OTHER INFORMATION: 5'-LTR
```

```
; NAME/KEY: 5'UTR
; LOCATION: (532)..(1219)
; FEATURE:
```

```
; NAME/KEY: mat_peptide
; LOCATION: (1220)..(5062)
; OTHER INFORMATION: m4 mdr-1 cDNA
```

```
; NAME/KEY: misc feature
; LOCATION: (5215)..(5774)
; OTHER INFORMATION: 3'-LTR
```

```
; NAME/KEY: misc feature
; LOCATION: (5775)..(8630)
; OTHER INFORMATION: plasmid backbone (pUC)
```

```
; NAME/KEY: misc feature
; LOCATION: (1)-(8630)
; OTHER INFORMATION: retroviral expression vector SFbeta71m4
```

Alignment Scores:

Pred. No.:	2,68e-157	Length:	8630
Score:	1570.00	Matches:	305
Percent Similarity:	76.85%	Conservative:	100
Best local Similarity:	57.87%	Mismatches:	100
Query Match:	60.57%	Indels:	22
DB:	10	Gaps:	5

US-09-873-409-8 (1-514) x US-09-306-417-1 (1-8630)

```
QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
DB 1370 ATGGTGGTGGGAACCTTTGGTGCATCCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 1429
QY 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
DB 1430 GTGTTGGAGAATGACATATCTTTGCCAAATGTCAGGAAATTTAGAAATCTGATGTCA 1489
QY 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
DB 1490 AACATCACTAATAGAGTGTATCATGATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 1549
QY 44 -----ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
DB 1550 ATGACCAGATATGCTTATTATACAGTGAATTTGGTGGGTGCTGGTTGCTGCTTAC 1609
QY 62 IleGlnIleSerLeuTrpIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
DB 1610 ATTCAGTTTCAATTTTGGTGGCAGCTGGGAAGCAAAATACAAAAATTAGAAAACAG 1669
QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101
DB 1670 TTTTTCATGCTATATATGACAGGAGATAGGCTGTTTGTATGTGCACCATGTGTGGGAG 1729
QY 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
```



```
Db 1730 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAAAGATTATTGGTGACAAAATT 1789
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db 1790 GGAATGTTCTTTTTCAGTCAATGGCAACATTTTTCACGGGTTTATAGTAGGATTACACGT 1849
Qy 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db 1850 GGTTGGAAGCTAACCCCTTGATTTTGGCCATCAGTCTCTTCTTGACACTCAGCTGCT 1909
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGlnLeuSerAlaThrSerLysAla 180
Db 1910 GTCTGGCAAGATGATCTATCTTCTTACTATAAGAACTCTTAGCGTATGCAAAAGCT 1969
Qy 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db 1970 GGAGCAGTACTGAAGAGGCTTGGCAGCAATTAGAATCTGTGATTCATTTGGAGACAA 2029
Qy 201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
Db 2030 AAGAAAGACTTGAAGGTACACAAAAATTTAGAAGAACTAAAGAAATTTGGATAAAG 2089
Qy 221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
Db 2090 AAAGCTATTACAGCAATAATTTCTATAGGTGCTGCTTCTCTGTGATCTATGATCTTAT 2149
Qy 241 GlyLeuAlaPheTyrTyrGlyThrSerLeuIleAsnGlyGluProGlyTyrThrIle 260
Db 2150 GCTCTGGCCCTTCTGGTATGGACACCTTGGTCTCTCAGGGGAA-----TATTTCTATT 2203
Qy 261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
Db 2204 GGACAGTACTCACTGATTTCTTCTGTATTAATTTGGGGCTTTTAGTGTGGACAGGCA 2263
Qy 281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
Db 2264 TCTCCAAGCATTTGAAGCATTTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 2323
Qy 301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
Db 2324 ATTGATTAATGAAGCAATTTGACAGCTATTTCGAAGATGGGCACAAACCAAGATAATATT 2383
Qy 321 GlyGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
Db 2384 AAGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCAATCTCGAAAGAAAGTTAAG 2443
Qy 341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
Db 2444 ATCTTGAAGGCGCTGAACCTGAAGGTGCAGAGTGGCAGACGGTGGCCCTGTTGGAAC 2503
Qy 361 AsnGlySerGlyLysSerThrValValGlnLeuGlnArgLeuTyrAspProAspAsp 380
Db 2504 AGTGGCTGTGGGAAGACACAACTGATCCAGTCCAGTGCAGAGGCTCTATGACCCCAAGAG 2563
Qy 381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
Db 2564 GGGATGTCAGTGTGTGATGGACAGGATATTAGGCCAATAATGTAAAGTTTCTACGGAA 2623
Qy 401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
Db 2624 ATCATTTGGTGTGGTGAAGTTCAGGAACCTGTATTGTTGCCACCAAGATAGCTGAAACATT 2683
Qy 421 LysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
Db 2684 CGCTATGGCGTGAATAATGTCCATCGATGAGATTGAGAAAGCTGCAAGGAAGCCCAAT 2743
Qy 441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
Db 2744 GCCTATGACTTTATCATGAAGAACTGCCTCATAAATTTGACACCTTGGTGGAGAGAGGG 2803
Qy 461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db 2804 GCCCAGTTGAGTGTGGGCGAAGACAGAGGATCGCCATTGACGTGCCCTGTTTCGCAAC 2863
```

```
Qy 481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db 2864 CCCAAGATCTCTCTGCTGATGAGGCCAGCTCAGCTTGGACACAGAAAGCGAAGCAGTG 2923
```

```
Qy 501 ValGlnAlaAlaLeuGluLys 507
Db 2924 GTTCAGGTGGCTCTGGATAAG 2944
```

RESULT 12

```
US-09-306-417-2
; Sequence 2, Application US/09306417
; Patent No. US20020103144A1
; GENERAL INFORMATION:
; APPLICANT: Heinrich-Pette-Institut
; TITLE OF INVENTION: Retroviral Gene Transfer Vectors
; FILE REFERENCE: P50491
; CURRENT APPLICATION NUMBER: US/09/306,417
; CURRENT FILING DATE: 1999-05-06
; EARLIER APPLICATION NUMBER: DE 198 22 115
; EARLIER FILING DATE: 1998-05-08
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 8630
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: proviral
; OTHER INFORMATION: plasmid DNA
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(8630)
; OTHER INFORMATION: retroviral expression vector Sfbeta91mSA1
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(160)
; OTHER INFORMATION: plasmid backbone (pUC)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (161)..(577)
; OTHER INFORMATION: 5'-LTR
; NAME/KEY: 5'UTR
; LOCATION: (532)..(1219)
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (1220)..(5062)
; OTHER INFORMATION: mSA1 mdr1 cDNA
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (5215)..(5774)
; OTHER INFORMATION: 3'-LTR
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (5775)..(8630)
; OTHER INFORMATION: plasmid backbone (pUC)
US-09-306-417-2
```

```
Alignment Scores:
Pred. No.: 2,68e-157 Length: 8630
Score: 1570.00 Matches: 305
Percent Similarity: 75.85% Conservative: 100
Best Local Similarity: 57.87% Mismatches: 100
Query Match: 60.57% Indels: 22
DB: 10 Gaps: 5
```

```
US-09-873-409-8 (1-514) x US-09-306-417-2 (1-8630)
```

```
Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 1370 ATGGTGGTGGGAACATTTGGCTGCCATCATCCATGGGGCTGGACTTCTCTCATGATGCTG 1429
```



```

Qy 21 ValLeuGlyGluMetSerHsp-----AsnLeuIleSer 31
Db 1430 GTGTTTGGAGAAATGACAGATATCTTTGCAAAATGACAGAAATTTAGAAATCTGATGCA 1489
Qy 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
Db 1490 AACATCACTAATGAAGTGAATATCAATGATACAGGTTCTTTCATGAATCTGGAGGAGAC 1549
Qy 44 -----ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db 1550 ATGACACAGGTATGCTTATTATACAGTGGAAATTTGGTGTGGGTGCTGTGTGCTTAC 1609
Qy 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db 1610 ATTCAGGTTTCATTTTGGTGTGGTGGCAGGAGATAGGCTGTTTGTGTCACGATTTAGAAACAG 1669
Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101
Db 1670 TTTTTCATGCTATATATGCGACAGGATAGGCTGTTTGTGTCACGATTTGGGGAG 1729
Qy 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db 1730 CTTAACACCCGACTTACAGATGATGCTCTAAAGATTAAATGAAGTTATTGTCGACAAAT 1789
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db 1790 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGGATTTACACGT 1849
Qy 141 GlyTyrLeuLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db 1850 GGTGGAAGCTAACCTTGTGATTGTTGGCCATCAGTCCCTGTTCTTGGACTGTCAGCTGT 1909
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db 1910 GTCTGGCAAGACTACTATCTTACTTACTGATAAGAACTCTTACGATGCAAAAGCT 1969
Qy 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db 1970 GGAGCAGTAGCTGAAGAGCTTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 2029
Qy 201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyLys 220
Db 2030 AAGAAGAACTTGAAGAGGTACAAACAAATTTAGAAAGAGCTTAAAGAAATGGGATAAG 2089
Qy 221 ArgThrIleAlaSerLysValSerLeuGlyValAlaValTyrPhePheMetAsnGlyThrTyr 240
Db 2090 AAAGCTATTACGCCAATATTCTATAGTGTGCTTCTCTGCTGATCTATGCACTTAT 2149
Qy 241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
Db 2150 GCTCTGGCCTTCTGGTATGGGACCACTTGGTCTCTCAGGGGAA-----TATTCATT 2203
Qy 261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
Db 2204 GGACAAAGTACTCACTGATTTCTTCTGATTAAATTTGGGGCTTTTAGTGTGGACAGGCA 2263
Qy 281 ValProHisPheGluThrPheAlaIleAlaArgGlyValAlaPheHisIlePheGlnVal 300
Db 2264 TCTCCAAAGCATTTGCAATTTGCAATGCAAGGAGCAGCTTATGCAATCTTCAGATA 2323
Qy 301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
Db 2324 ATTGATAAAGCAAGTATTGACAGCTATTCGAAGAGTGGGCAACCAACAGATATAAT 2383
Qy 321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
Db 2384 AAGGGAAATTTGGAAATTCAGAAATGTTCACTTCAGTTACCCCATCTCGAAAGAGTTAAG 2443
Qy 341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
Db 2444 ATCTGAAGGGCTGAACCTGAAGGTGCAAGTGGCAGCAGCGTGGCTTGGTTGAAAC 2503
Qy 361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380

```

```

Db 2504 AGTGCTGTGGAAAGACACACAGTCAGCTGATGAGAGCTCTATGACCCACAGAG 2563
Qy 381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
Db 2564 GGGATGGTCAGTGTGTGATGGACAGATATTAGGACCATAAATGTAAGGTTTCTACGGAA 2623
Qy 401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnIle 420
Db 2624 ATCATTTGGTGTGTGAGTCAGGAACCTGTATTGTTTGGCACCACCATAGCTGAAACATT 2683
Qy 421 LysTyrGlyArgAspValThrAspGluMetGluArgAlaAlaArgGluAlaArg 440
Db 2684 CGCTATGGCCGGAATGTACCATGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 2743
Qy 441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGly 460
Db 2744 GCCTATGACTTATCATGAAACTGCTCATAAATTTGACACCTTGTGGAGAGAGGG 2803
Qy 461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db 2804 GCCCAGTTGAGTGTGGGAGAGAGGATCGCATTTGCACTGCTGCTGTTGCGCAAC 2863
Qy 481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db 2864 CCCAAGATCCTCTCTGATGAGGCCAGTCAGCTTGGACACAGAAAGCGAAGCATG 2923
Qy 501 ValGlnAlaAlaLeuGluLys 507
Db 2924 GTTCAGTGGCTCTGGATAAG 2944

```

RESULT 13

```

US-09-769-097-3
; Sequence 3, Application US/09769097
; Patent No. US20020055128A1
; GENERAL INFORMATION:
; APPLICANT: Kimberly Anne Brun
; APPLICANT: Richard James Chenery
; APPLICANT: Harma Ellens
; APPLICANT: John Anthony Feild
; APPLICANT: Lin Yue
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES
; TITLE OF INVENTION: ENCODING RAT MDRIA AND SCREENING METHODS THEREOF
; FILE REFERENCE: GP-50009-C2
; CURRENT APPLICATION NUMBER: US/09/769,097
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 09/208,809
; PRIOR FILING DATE: 1998-12-09
; PRIOR APPLICATION NUMBER: 09/156,800
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: US99/20770
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 4425
; TYPE: DNA
; ORGANISM: RATTUS RATTUS
US-09-769-097-3

```

Alignment Scores:

```

Pred. No.: 2,3e-157 Length: 4425
Score: 1566.50 Matches: 300
Percent Similarity: 76.82% Conservative: 101
Best Local Similarity: 57.47% Mismatches: 102
Query Match: 60.44% Indels: 19
DB: 10 Gaps: 4

```

US-09-873-409-8 (1-514) x US-09-769-097-3 (1-4425)

```

Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 499 ATGCTGTGGAACTCTGGCGGCCATTATCCATGGAAATTCGCTCCCACTTATGATGCTG 558

```


Db 499 ATGCTGCTGGGAACCTCTGCGGCCATTATCCATGGAAATGCGCTCCCATCTTATGATGCTG 558
 Qy 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
 Db 559 GTCTTTGGGAGACATGACAGTCTTTGCAAT- - - -GTAGGAACAACACCGTAGTATG 612
 Qy 41 SerPhePheArgLeuThr- 46
 Db 613 AGTTTCTACAATGCTACAGACATATATGCAAGCTGGAGGACGAATGGCCACGTACGCC 672
 Qy 47 LeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeu 66
 Db 673 TACTATTACAGCGCAATGCTGCGGTGTCTCATCTGCTCCATCATCCAGGTTTCACIT 732
 Qy 67 TrpIleIleThrAlaAlaArgGlnThrLysArgIleAryLysGlnPhePheHisSerVal 86
 Db 733 TGGTGCCTGGCAGCTGGGAGACAAATACAAAGATTAGGCAGAAAGTTTTTCCATGCCATC 792
 Qy 87 LeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGluLeuAsnThrArgMet 106
 Db 793 ATGAATCAGGAGATAGCTGCTTGTGACGTGCATGACGTTGGGAGCTCAACACCCGCGCTC 852
 Qy 107 Thr- - - - -AspIleAspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGln 125
 Db 853 ACAGATGACGCTCCAAAATTAATGAAGAAATTTGGTGCACAAATTTGGAATGTTCTTTTCAG 912
 Qy 126 AsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLysGlyTrpLysLeuThr 145
 Db 913 GCATGGCAACATTTTGGTGGTTTTTATATAGGATTTACTCGCGCTGGAAAGCTTAAT 972
 Qy 146 LeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAlaAlaCysSerArgMet 165
 Db 973 CTGTGATTTGGCCATCAGCCCTGTTCTTGGACTGTCAGCTGTTATTTGGGCAAGATA 1032
 Qy 166 ValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGlu 185
 Db 1033 TTGTCTTCATTTACTGATGAAGAACTCCAGGCTTATGCAAAAGCTGGAGCAGTTGCTGAA 1092
 Qy 186 GluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGln 205
 Db 1093 GAAGTCTTAGCAGCCATCAGACTGTGATGCTTTGGAGGACAAAAGAAAGAACTTGAA 1152
 Qy 206 ArgTyrThrGlnAsnLeuLysAlaLysAspPheGlyIleLysArgThrIleAlaSer 225
 Db 1153 AGGTACAATAACAATTTGAAGAAGCTTAAAGCTTGGGATAAGAAGACTATACCGGC 1212
 Qy 226 LysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrp 245
 Db 1213 AACATTTCCATGGTGCAGCTTTTCTGCTTATCTATGTCATCATATCTCTGGCATTTCTG 1272
 Qy 246 TyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAla 265
 Db 1273 TATGGACTTCTTGTGCTATCTCAAAAGAA- - - -TACACTATGGACNAGTGCTCACT 1326
 Qy 266 ValPhePheSerValIleHisSerTyrCysIleGlyValAlaAlaValProHisPheGlu 285
 Db 1327 GTCTTTTCTGTATTAATTTGGAGCATTCAGGTGGGAGGAGCATTCACAAATATTGAA 1386
 Qy 286 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPro 305
 Db 1387 GCCTTCGCCAATGCTAGAGGAGCAGCTTATGAAGTCTTTCAGTATTAATTAATAAGCCC 1446
 Qy 306 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 325
 Db 1447 AGTATAGACGCTTCTCAAGAGTGGGACCAAAACCCGACAACTACAAAGAAATTTGGAA 1506
 Qy 326 PheLysAsnValSerPheAntYrProSerArgProSerIleLysIleLeuLysGlyLeu 345
 Db 1507 TTCAAAATAATTACATTCACTTCACTTACCTGCTCGAAAGACGTTTCAGATCTTTGAAGGCCCTC 1566
 Qy 346 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 365
 Db 1567 AACCTGAAGGTGAAGAGCGGACAGCGTAGCCCTGTTGGCAACAGCTGCTGTTGGGAAA 1626

Qy 366 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 385
 Db 1627 AGCAACAACCTGTCCAGCTGTCAGAGGCTTACGACCCCATAGAGGGCGAGGTCAGTATC 1686
 Qy 386 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal 405
 Db 1687 GACGGACAGGACATCAGGACCATCAATGTAGGTATCTCGGGAAATCAATGGGGTGGTG 1746
 Qy 406 SerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAsp 425
 Db 1747 ACTCAGAACCCGCTGCTGTTTCCACCAACAATTCGCGAAACATTCGCTATGCGCGAGAA 1806
 Qy 426 AspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIle 445
 Db 1807 AACGTCACCATCGATGAGTAAAGCTGTCAAGGAAGCAATGCTATGATTTTCATC 1866
 Qy 446 MetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyValAlaGlnMetSerGly 465
 Db 1867 ATGAACCTGCCCAACAATTTGACACCTCTGTTGGTGGAGAGGGCGCAGCTGAGTGGG 1926
 Qy 466 GlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIle 485
 Db 1927 GGACAGAAACAGAGGATGCCATTTGCCGGGCGCTTGGTCCGCAACCCCAAGATCCTTTG 1986
 Qy 486 LeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaLeu 505
 Db 1987 TTTGATGAGGCCACGCTGACCTTGGACACAGAAAGCGAGCGGTGTTCAGGCGGCTCTG 2046
 Qy 506 GluLys 507
 Db 2047 GATAAG 2052
 RESULT 15
 US-10-072-621-2
 ; Sequence 2, Application US/10072621
 ; Patent No. US20020169137A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Reiner, Peter B.
 ; APPLICANT: Connop, Bruce P.
 ; APPLICANT: Pollard, Michelle
 ; TITLE OF INVENTION: REGULATION OF AMYLOID PRECURSOR PROTEIN EXPRESSION
 ; TITLE OF INVENTION: BY MODIFICATION OF ABC TRANSPORTER EXPRESSION OR ACTIVITY
 ; FILE REFERENCE: 100103.402
 ; CURRENT APPLICATION NUMBER: US/10/072.621
 ; CURRENT FILING DATE: 2002-02-08
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 4643
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-072-621-2
 Alignment Scores:
 Pred. No.: 6,64e-157 Length: 4643
 Score: 1562.50 Matches: 305
 Percent Similarity: 76.85% Conservative: 100
 Best Local Similarity: 57.87% Mismatches: 99
 Query Match: 60.28% Indels: 23
 DB: 9 Gaps: 6
 US-09-873-409-8 (1-514) x US-10-072-621-2 (1-4643)
 Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
 Db 575 ATGTGTGTGGGAACCTTGTGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 634
 Qy 21 ValLeuGlyGluMetSerAsp- 31
 Db 635 GTGTTGGAGAAATGACAGTATCTTTGCAATGCAAGAAATTTAGAACATCTGATGTCA 694
 Qy 32 GlyCysLeuValGlnThr- 43


```
Db 695 AACATCACTAATAGAGTGTATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 754
      ::::: ||| |||||
Qy 44 -----ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 755 ATGACCAGGTATGCCATTATTATACAGTGAATTTGGTGGGTGCTGGTCTTAC 814
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 815 ATTACAGTTTCATTTTGGTCCCTGGGAGCTGGAGACAAATATACACAAATTTAGAAACAG 874
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 875 TTTTTCATGCTATAATGCACAGGAGATAGGCTGGTTGATGTGCACGATGTGGGAG 934
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 102 LeuAsnThrArgMetThr----AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 935 CTTAAACACCCGACTTACAGATGATGTCTCCAAAGATTAAATGAAGGAATTTGGTGACAAAT 994
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 995 GGAATGTTCTTTTTCAGTCATGGAACAAATTTTCTACCTGGGTTTATAGTAGGATTACAGT 1054
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1055 GGTGGAAGCTAACCCCTTGATTTTGGCCATCAGTCTCTGTTCTGGAGCTGTCAGCTGCT 1114
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGlnLeuSerAlaTyrSerLysAla 180
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1115 GTCTGGGCAAGATACCTATCTTCTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 1174
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 181 GlyValAlaAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1175 GGAGCAGTACTGAAGAGGCTTGGCAGCAATTAGAACTGTGATTCATTTGGAGGACAA 1234
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1235 AAGAAAGAACTTGAAGGTACACAAAAATTTAGAGAAGCTAAAGAAATTTGGATAAAG 1294
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPheMetAsnGlyThrTyr 240
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1295 AAAGCTATTACAGCCAAATTTCTATAGGTGCTGCTTTCTGCTGATCTATGCACTCTAT 1354
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleAsnGlyGluProGlyTyrThrIle 260
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1355 GCTCTGGCCCTTCTGGTATGGACCACTTGTCTCTCAGGGAA-----TATTCTATT 1408
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1409 GGACAACTACTCACTGTA---TTTTCTGTATTAATTTGGGCTTTTAGTGTGCACAGGCA 1465
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1466 TCTCCAAGCATTTGAAGCATTTGCAAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1525
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1526 ATTGATAAATGAAGCAATTTGAGCTATTTCGAGGTGGGACAAACCCAGATATATT 1585
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1586 AAGGGAATTTGGAATTCAGAAATGTTCACTTCAGTTACCACTCGAAGAAAGTTAAG 1645
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1646 ATCTTGAAGGCTCTGAACCTCTGAAGTGCAGAGTGGGACAGCGTGGCCCTGGTTGGAAC 1705
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnLeuTyrAspProAspAsp 380
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1706 AGTGGCTGTGGGAAGAGACAAACAGTCCAGCTGATGAGAGGCTCTATGACCCACAGAG 1765
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Search completed: March 31, 2003, 16:15:25
Job time : 246.092 secs

```
Db 1766 GGGATGGTCAAGTGTTCATGGACAGGATATTAGGACCATAAATGTAAAGTTTCTACGGGAA 1825
Qy 401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1826 ATCAATTGGTGTGGTGAAGTCTGATTTGTTTCCACCACGATAGCTGAAAACATT 1885
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 421 LysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1886 CGCTATGGCCGTGAAAATGTCAACCATGGATGATGAGAAAGCTGTCAAGGAAGCCAAT 1945
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGly 460
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1946 GCCTATGACTTTTATCATGAACCTGCCTCATAAATTTGCACCCCTGGTTGGAGAGAGGG 2005
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2006 GCCAGTTGAGTGGTGGGAGAGCAGAGGATCGCCATTTGCACGTGCCCTGGTTCCGCAAC 2065
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2065 CCCAAGATCCTCTGCTGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 2125
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 501 ValGlnAlaAlaLeuGluLys 507
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2126 GTTCAGGTGGCTCTGGATAAG 2146
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```


GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 30, 2003, 01:24:25 ; Search time 6948.72 Seconds
(without alignments)
2152.747 Million cell updates/sec

Title: US-09-873-409-8
Perfect score: 2592
Sequence: 1 MILGILASLVNGACILPLMLP.....SEKSAVQAALKEKTPRYSF 514

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -MODEL=frame+ p2n model -DEV=xlh

-Q/cgn2_1/USFTO_spool/US09873409/runat_27032003.115418.19228/app_query.fasta_1.7544
-DB=GenEmbl -OFT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09873409 @CGN 1.1 54887 @runat_27032003.115418.19228 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl :
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sv.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*

29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rodi.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sv.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2592	100.0	1940	6	AX339034 Sequence
2	2568.5	99.1	2021	6	AX339033 Sequence
3	2552	98.5	3621	6	AX339032 Sequence
4	2528.5	97.6	3702	6	AX339031 Sequence
5	2460.5	94.9	3699	6	AX478104 Sequence
6	1699.5	65.6	3177	6	AX339030 Sequence
7	1590	61.3	4189	6	AX322791 Sequence
8	1590	61.3	4298	10	MUSMDR
9	1585	61.1	4279	6	AX105082 Sequence
10	1581	61.0	4279	6	AX105078 Sequence
11	1581	61.0	4279	6	AX105080 Sequence
12	1580.5	61.0	4195	6	AX108656 Sequence
13	1580	61.0	4045	12	AF269224 Synthetic
14	1579	60.9	3860	6	AX322787 Sequence
15	1579	60.9	4378	6	E02326
16	1579	60.9	4646	6	AX391099 Sequence
17	1578.5	60.9	4280	10	CRUPGPII
18	1576	60.8	3888	10	AY082609
19	1576	60.8	4233	6	AR123273 Sequence
20	1576	60.8	4279	6	AX105057 Sequence
21	1576	60.8	4317	4	AF045016 Canis fam
22	1576	60.8	4317	6	AX105059 Sequence
23	1576	60.8	4669	6	AR091275 Sequence
24	1576	60.8	4669	6	AR203322 Sequence
25	1572	60.6	4186	6	AX108654 Sequence
26	1571.5	60.6	4254	6	AX401748 Sequence
27	1571.5	60.6	4927	10	RATMDRM
28	1571.5	60.6	4927	10	AF257746
29	1570.5	60.6	4323	10	AF286167
30	1570	60.6	3860	6	AX322789 Sequence
31	1570	60.6	4646	6	AX336420 Sequence
32	1570	60.6	4646	6	AX336708 Sequence
33	1570	60.6	4646	6	I49610
34	1570	60.6	4646	9	HUMMDR1
35	1570	60.6	4669	6	I08557
36	1570	60.6	6505	6	AR028671 Sequence
37	1570	60.6	8630	6	AX012320 Sequence
38	1570	60.6	8630	6	AX012321 Sequence
39	1570	60.6	9318	6	AR028672 Sequence
40	1568.5	60.5	4264	6	AR051647 Sequence
41	1568.5	60.5	4264	6	AR051650 Sequence
42	1565.5	60.4	4356	10	MUSMDRAA
43	1565	60.4	4296	10	CRUPGPI
44	1564	60.3	4304	10	CRUPGPI165
45	1563.5	60.3	3489	4	AB029153 Felis cat

ALIGNMENTS

AX339034
LOCUS AX339034 1940 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 16 from Patent WO0194400.
ACCESSION AX339034
VERSION AX339034.1 GI:18129126
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE Frank, M.H. and Sayegh, M.H.
AUTHORS
TITLE A gene encoding a multidrug resistance human p-glycoprotein
JOURNAL homologue on chromosome 7p15-21 and uses thereof
Patent: WO 0194400-A 16 13-DEC-2001.
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES
Location/Qualifiers
source 1. 1940
/organism="Homo sapiens"
/db xref="taxon:9606"
BASE COUNT 530 a 394 c 448 g 568 t
ORIGIN
Alignment Scores:
Pred. No.: 2.68e-192 Length: 1940
Score: 2592.00 Matches: 514
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-873-409-8 (1-514) x AX339034 (1-1940)
Qy 1 MetIleuGlyIleLeuAlaSerIleuValIasnGlyAlaCysLeuProLeuMetProLeu 20
Db 34 ATGATCTGGGTATCTGGCATCTGGTCAATGGAGCCTGCCTTCCTTAATGCCACTG 93
Qy 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
Db 94 GTTTTAGAGAAATAGTGATTAACCTTATTAGTGGATGTCTAGTCCAACTAACACATAC 153
Qy 41 SerPheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGly 60
Db 154 TCTTTCTTCAGGTTGACCTGTATTATGTTGAATAGTGTGCTGCTTCAATTTTGGT 213
Qy 61 TyrIleGlnIleSerLeuTrpIleIleThrAlaIaargGlnThrIlysArgIleArgIys 80
Db 214 TACATACAGATTTCCTTGTGGATTATTAACTCAGCAGCAGACACCAAGAGGATTTCGAAA 273
Qy 81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGly 100
Db 274 CAGTTTTTTCATTCAGTTTTGGCACAGGACATCGCTGGTTTGTAGTGTGACATCGGT 333
Qy 101 GluLeuAsnThrArgMetThrAspIleAspIlysIleSerAspGlyIleGlyAspIysIle 120
Db 334 GAACCTTAACACTCGCATGCAGACATTCGACAAATCAGTGATGTTATTGGAGATAAGATT 393
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValIys 140
Db 394 GCTCTGTTGTTTCAAAACATGCTACTTTTTTCGATTGGCCCTGGCAGTTGGTTTGGTGAAG 453
Qy 141 GlyTrpIysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db 454 GGCTGGAAACTCACCTAGTGACTCTATCCAGTCTCTCTTATTAATGGCTTCAGCGGCA 513
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerIysGluLeuSerAlaTyrSerIysAla 180
Db 514 GCATGTTCTAGGATGGTTCATCTCATTTGACAGGATTAAGGAATTAAGTCTATTTCGAAAGCT 573
Qy 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db 574 GGGGCTGTGGCAGAGAAGTCTTGTTCATCAATCCGAACAGTCATAGCCTTTAGGGCCCG 633

Qy 201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
Db 634 GAGAAAGAACTTCAAAGGTATACACAGAATCTCAAAGATGCAAGAGGATTTCGCATAAAA 693
Qy 221 ArgThrIleAlaSerIysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
Db 694 AGGACTATAGCTTCAAAGAGTCTCTTGGTGTGTGTACTTCTTTATGAATGAACCTAT 753
Qy 241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
Db 754 GGACTTGTCTTTTGGTATGGAACCTCTTGATTTCTTAATGGAGAACCTGGATATACCATC 813
Qy 261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
Db 814 GGGACTGTTCTTCTGCTGTTCTTTCTTTAGTGAATCCATAGCAGATTATTGCAATGGAGCAGCA 873
Qy 281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
Db 874 GTCCTCTACCTTTGAAACCTTCGCAATAGCCGAGGAGCTGCCTTTTCATATTTTCCAGGTT 933
Qy 301 IleAspIysIysProSerIleAspAsnPheSerThrAlaGlyTyrIysProGluSerIle 320
Db 934 ATTGATAAGAAACCCAGTATAGATAAATCTTTCCACAGCTGGATATAAACCTGAATCCATA 993
Qy 321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
Db 994 GAAGGAACCTGTGGAATTTAAATAATGTTCTTTCAATATCATCAAGACCATCTATCAAG 1053
Qy 341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
Db 1054 ATTCTGAAGGTCTGAATCTCAGAAATTAAGTCTGGAGAGACAGTCGCTTGGTGGGTCTC 1113
Qy 361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
Db 1114 AATGGCAGTGGGAGAGTACGGTAGTCCAGCTTCTCGAGAGGTTATATATCCGGATGAT 1173
Qy 381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
Db 1174 GGCTTTTATCATGTGTGATGAGATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGAC 1233
Qy 401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
Db 1234 CATATTGGAGTGTGTAGTCAAGAGCCCTGTTTGTTCGGGACCACCATCAGTAACAATATC 1293
Qy 421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
Db 1294 ARGATGGACGAGATGATGTGACTGATGAAGAGATGGAGAGACAGCAGGAGCAAAAT 1353
Qy 441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
Db 1354 GCGTATGATTTTATCATGTGAGTTTCTCTAATAATTTAATAATTTAGTGGGAAAAAGGA 1413
Qy 461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db 1414 GCTCAATAGTGGAGGGCAGAAACAGAGATCGCAATTCCTGCTGCTTGTAGTTCGAAC 1473
Qy 481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db 1474 CCCAAGATTCTGATTTTAGATGAGGCTACGCTGCCCCCTGGATTTCAGAAAGCAAGTCAGCT 1533
Qy 501 ValGlnAlaAlaLeuGluLysAspThrProArgTyrSerPhe 514
Db 1534 GTTCAAGCTGCACCTGGAGAGGATACCCCCAGGTATTCATTT 1575

RESULT 2

AX339033
LOCUS AX339033 2021 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 15 from Patent WO0194400.
ACCESSION AX339033
VERSION AX339033.1 GI:18129125
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 Frank, M.H. and Sayegh, M.H.

A gene encoding a multidrug resistance human p-glycoprotein

homologue on chromosome 7p15-21 and uses thereof

Patent: WO 0194400-A 15 13-DEC-2001;

THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)

Location/Qualifiers

1. .2021

/organism="Homo sapiens"

/db_xref="taxon:9606"

BASE COUNT 547 a 406 c 462 g 605 t 1 others

ORIGIN

Alignment Scores:

Pred. No.: 1-89e-190 Length: 2021
Score: 2568.50 Matches: 514
Percent Similarity: 95.01% Conservative: 0
Best Local Similarity: 95.01% Mismatches: 0
Query Match: 99.09% Indels: 27
DB: 6 Gaps: 1

US-09-873-409-8 (1-514) x AX339033 (1-2021)

```
QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
DB 34 ATGATCTGGGTATCTGGCATCACTGGTCAATGGAGCGCTTCCTTTAAATGCCACTG 93
QY 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
DB 94 GTTTTATAGGAGAAATGAGTGATACCTATTATAGTGGATGTCTAGTCCAAACTAACACATC 153
QY 41 SerPhePheArgLeuThrLeuTyrValGlyIleGlyValAlaLeuIlePheGly 60
DB 154 TCCTTTCTCAGGTGACCCGTGATATATGTTGGAATAGTGTCTGCTGCTTGTATTTTGGT 213
QY 61 TyrIleGlnIleSerLeuThrIleIleThrAlaAlaArgGlnThrLysArgIleArgLys 80
DB 214 TACATACAGATTTCCTTGTGGATTATAAATGCAGCAGCAGACCAAGAGATTCCGAAAA 273
QY 81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGly 100
DB 274 CAGTTTTTTTCATTTCAGTTTGGCAGAGACATCGGTGGTTGATAGCTGTGACATCGGT 333
QY 101 GluLeuAsnThrArgMetThrAspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
DB 334 GAACCTTAACACTCGCATGACAGACATTCGAAATATCAGTGTATGTTGGAGATAAGATT 393
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
DB 394 GCTCTGTTGTTTCAAAACATGTCTACTTTTTCGATTGGCCTGGCAGTTGTTGGTGAAG 453
QY 141 GlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
DB 454 GGTGGAACCTCACCTAGTACTCTATCCACGCTCTCCTCTTATAATGGCTTCAGCGGCA 513
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGlnLeuSerAlaTyrSerLysAla 180
DB 514 GCATGTTCTAGGATGGTCATCTCATGACCAGTAAGGAATTAAGTGCTTATTCGAAGCT 573
QY 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
DB 574 GGGGCTGTGGCAGAGAAGCTTGTGTCATCAATCCGAACAGTCATAGCTTTTAGGCCCCAG 633
QY 201 GluLysGluLeuGln----- 205
DB 634 GAGAAAGAACTTCAAGAGTCTTTCTTTTAAATATAACAAGATATGCTGGTTTTATTTT 693
QY 206 -----ArgTyrThrGlnAsnLeuLysAsp 213
DB 694 CCCAGTGGCTACTAAGTTGTTCTGTTNTTGTAAAGGTATACACAGAAATCTCAAGAT 753
```

```
QY 214 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 233
DB 754 GCAAGAGATTTTGGCATAAAAAAGGACTATAGCTTCAAAAGTGTCTCTGGTGTGTAC 813
QY 234 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsn 253
DB 814 TTTCTTTATGAAATGGAACCTATGGACTTGTCTTTTGGTATGGAACCTCCTTGATCTTAAT 873
QY 254 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 273
DB 874 GGAGAACTGGATATACCATCGGACTGTCTTCTGCTGTTTCTTTAGTGTAAATCCATAGC 933
QY 274 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 293
DB 934 AGTTATTGCAATTGGAGCAGCAGTCCCTCACTTTGAAACCTTCGCAATAGCCGAGAGCT 993
QY 294 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 313
DB 994 GCCTTTTCATATTTTCCAGGTTATTTGATAAGAAACCCAGTATAGATAACTTTTCCAGCT 1053
QY 314 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 333
DB 1054 GGATATAAACCTGAATCCATAGAGGAACCTGTGGAATTTAAAAATGTTCTTTCAATTAT 1113
QY 334 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 353
DB 1114 CCATCAAGACCATCTATCAAGATTCTGAAAGGTCTGAAATCTCAGAAATTAAGTCGAGAG 1173
QY 354 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 373
DB 1174 ACAGTCGCTTGTGCTGCTCAATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAG 1233
QY 374 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 393
DB 1234 AGTTATATGATCCGATCGATGCTTATCATGGTGGTGAAGATGCATCAGAGCTTTA 1293
QY 394 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 413
DB 1294 AATGTCGCGCATTTATCGAGACCATATTGGAGTGGTGTAGTCAAGAGCGCTGTTTGTGCGG 1353
QY 414 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 433
DB 1354 ACCACCATCAGTAACAAATATCAAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAG 1413
QY 434 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 453
DB 1414 AGAGCAGCAGGAGGAGCAATGGTATGATTTATCATGGAGTTTCTTAATAATTTAAT 1473
QY 454 ThrLeuValGlyLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 473
DB 1474 ACATTGTTAGGGGAAAAAGGAGCTCAAAATGAGTGGAGGGCAGAAACAGAGGATCGCAATT 1533
QY 474 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 493
DB 1534 GCTCGTGGCTTGTAGTTCGAAACCCCAAGATTCTGATTTTATAGTGGAGCTACGCTGCGCTG 1593
QY 494 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAspThrProArgTyrSer 513
DB 1594 GATTTCAGAAAGCAAGTCAGCTGTTTCAAGCTGCACTGGAGAGGATACCCCGAGTATTCA 1653
QY 514 Phe 514
DB 1654 TTT 1656
```

RESULT 3

AX339032

LOCUS

DEFINITION Sequence 14 from Patent WO0194400.

ACCESSION AX339032

VERSION AX339032.1 GI:18129124

KEYWORDS

SOURCE human.

ORGANISM Homo sapiens

linear

DNA

3621 bp

AX339032

Sequence 14 from Patent WO0194400.

PAT 09-JAN-2002

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 Frank, M.H. and Sayegh, M.H.

A gene encoding a multidrug resistance human p-glycoprotein
homologue on chromosome 7p15-21 and uses thereof

Patent: WO 0194400-A 14 13-DEC-2001;

THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)

FEATURES

Location/Qualifiers

1..3621

/organism="Homo sapiens"

/db_xref="taxon:9606"

BASE COUNT 1081 a _696 c 842 g 1002 t

ORIGIN

Alignment Scores:

Pred. No.: 7.4e-189 Length: 3621
Score: 2552.00 Matches: 507
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.46% Indels: 0
DB: 6 Gaps: 0

US-09-873-409-8 (1-514) x AX339032 (1-3621)

Qy 1 MetileLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 34 ATGATCTGGGTATACCTGGCTACTGGTCAATGAGCGCTGCTTCTTTAATGCCACTG 93
Qy 21 ValLeuGlyGluMetSerAspAsnLeuSerGlyCysLeuValGlnThrAsnThrTyr 40
Db 94 GTTTTAGGAGAAATGAGTAACTTATTAGTGGATGCTAGTCCAAACTAACACATAC 153
Qy 41 SerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaLeuIlePheGly 60
Db 154 TCTTTCTTCAGGTTGACCTGTATTATGTTGGAATAGTGTGTGCTGCTGATTTTGGT 213
Qy 61 TyrlleGlnIleSerLeuTrpIleleThrAlaAlaArgGlnThrLysArgIleArgLys 80
Db 214 TACATACAGATTCTCTGTGGATATTAACTGACACACACAGACCAAGAGATTGAAAA 273
Qy 81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGly 100
Db 274 CAGTTTTTTCATTTCAGTTTGGCACAGGACATCGCTGGTTTGATGCTGTGACATCGT 333
Qy 101 GluLeuAsnThrArgMetThrAspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db 334 GAACCTTAACACTCGCATGACAGACATTTGACAAATCAAGTATGATGTTGGAGATAAG 393
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db 394 GCTCTGTTGTTTCAAAACATGCTACTTTTTCGATTTGGCTGGCAGTTGGTTGGTGA 453
Qy 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db 454 GGCTGGAAACTCACCTAGTACCTATCCAGGCTCTCTCTTATAATGGCTTCAGCGGCA 513
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db 514 GCATGTTCTAGGATGGTCACTCAATGACCAAGTAAGGAAATTAAGTGCCTATTCCAA 573
Qy 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db 574 GGGGCTGTGGCAGAAAGAGTCTTGTTCATCAATCCAGACATAGAGCTTTTAGGGCC 633
Qy 201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
Db 634 GAGAAAGAACTTCAAGGTATACACAGAACTCTCAAGATGCAAGAGATTGGCATAAA 693
Qy 221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
Db 694 AGGACTATAGCTTCAAAAGTGTCTCTGTTGCTGTGTACTTCTTTTATGAATGGAAC 753

Qy 241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
Db 754 GGACTTGCCTTTTGGTATGGAACCTCTTGAATCTTAAATGGAGAACCTGGATATACCATC 813
Qy 261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
Db 814 GGACACTGTTCTTCTGCTGTTTCTTTAGTGAATCAATGAGCAGTTATTGTCATTTGAGCAGCA 873
Qy 281 ValProHisPheGluThrPheAlaIleAlaArgGlyValAlaPheHisIlePheGlnVal 300
Db 874 GTCCCTCACTTTGAAACCTTCGCAATAGCCGAGGAGCTGCCCTTTTCAATATTTTCCAGGTT 933
Qy 301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
Db 934 ATTGATAAGBAACCCAGTAGATATACTTTCCACAGCTGGATATAAACCTGAATCCATA 993
Qy 321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
Db 994 GAAGGAACCTGTGGAATTTAAAAATGTTCTTTCAATATTCATCAAGACCATCTATCAAG 1053
Qy 341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
Db 1054 ATTCTGAAAGGCTCTGAATCTCAGAAATTAAGTCTGGAGACACATCGCCTTGGTGGTCTC 1113
Qy 361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
Db 1114 AATGGCAGTGGGAAGAGTAGCGTAGTCCAGCTTCTGCAGAGGTTATATGATCCGGATGAT 1173
Qy 381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
Db 1174 GGCTTTATCATGTGTGATGAGATGACATCAGAGCTTTTAAATGTGGGCATTTATCGAGAC 1233
Qy 401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
Db 1234 CATATTGGAGTGTGTAGTCAAGAGCCTGTTTGTTCGGGACCACTAGTAAACAATATC 1293
Qy 421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
Db 1294 AAGTATGGACGACATGATGTGACTGATGAAGAGATGGAGAGACAGCAAGGGAAGCAAT 1353
Qy 441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
Db 1354 GCCTATGATTTTATCATGAGTTCCTTAATAATTTAAATACATTGGTAGGGGAAAAAGGA 1413
Qy 461 AlaGluMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db 1414 GCTCAATGATGGAGGGCAGAAACAGAGATCGCAATGCTCGTCTTAGTTCGAAAC 1473
Qy 481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db 1474 CCCAAGATTCGATTTAGATGAGGCTACGTCCTGCTGGATTTCAGAAAGCAAGTCAGCT 1533
Qy 501 ValGlnAlaAlaLeuGluLys 507
Db 1534 GTTCAAGCTGCACTGGAGAG 1554

RESULT 4

AX339031

Sequence 13 from Patent WO0194400.

DEFINITION

AX339031

VERSION

AX339031.1 GI:18129123

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1

AUTHORS

TITLE

JOURNAL

linear PAT 09-JAN-2002

Frank, M.H. and Sayegh, M.H.
A gene encoding a multidrug resistance human p-glycoprotein
homologue on chromosome 7p15-21 and uses thereof
Patent: WO 0194400-A 13 13-DEC-2001;
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)

FEATURES		Location/Qualifiers	
source	1..3702	/organism="Homo sapiens"	
		/db_xref="taxon:9606"	
BASE COUNT	1098 a 708 c 856 g 1039 t	1 others	
ORIGIN			
Alignment Scores:			
Pred. No.:	5.1e-187	Length:	3702
Score:	2528.50	Matches:	507
Percent Similarity:	94.94%	Conservative:	0
Best Local Similarity:	94.94%	Mismatches:	0
Query Match:	97.55%	Indels:	27
DB:	6	Gaps:	1
US-09-873-409-8 (1-514) x AX339031 (1-3702)			
QY	1	MetIleuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu	20
Db	34	ATGATCTGGGTATCTGGCATCACTGGTCAATGGAGCCTGCCTCTCTTTAATGCCACTG	93
QY	21	ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr	40
Db	94	GTTTTAGAGAAATGAGTGATACCTTATTAGTGGATGCTAGTCCAACTAACACATAC	153
QY	41	SerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGly	60
Db	154	TCITTTCTCAGGTGGTACCTGTATTATGTTGGAATAGGTGTGCTGCCITGATTTTGGT	213
QY	61	TyrIleGlnIleSerLeuThrIleIleThrAlaAlaArgGlnThrIlysArgIleArgLys	80
Db	214	TACATACAGATTTCTCTGTGGATTATAACTGCAGCACGACAGACCAAGAGGATTCGAAA	273
QY	81	GlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGly	100
Db	274	CAGTTTTTCATTCAGTTTGGCAGAGCATCGCTGGTTTGTAGTGTGATCGATCGGT	333
QY	101	GluLeuAsnThrArgMetThrAspIleAspLysIleSerAspGlyIleGlyAspLysIle	120
Db	334	GAACCTTAACACTCGCATGCAGACATTCACAAATCAGTGATGGAGATAAGATT	393
QY	121	AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys	140
Db	394	GCTCTGTGTGTTTCAAAACATGTCTACTTTTTTCGATTGGCCTGGCAGTTGGTGAAG	453
QY	141	GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla	160
Db	454	GGCTGGAACCTCACCTAGTGACTCTATCCACGTCTCCTCTTATAATGGCTTCAGCGCA	513
QY	161	AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla	180
Db	514	GCATGTTCTAGGATGGTCACTCATTCAGCCAGTAAGGAATTAAGTGCCTATTCCAAAGCT	573
QY	181	GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln	200
Db	574	GGGGCTGTGGCAGAGAAGTCTTGTCATCAATCCGAACAGTCATAGCCTTTAGGGCCAG	633
QY	201	GluLysGluLeuGln-----	205
Db	634	GAGAAAGAACTTCAAAAGGTCTTTCCTTTTAAATATAACAAGATATGCTGGTTTATTTT	693
QY	206	-----ArgTyrThrGlnAsnLeuLysAsp	213
Db	694	CCCCAGTGGCTACTAAGTTGTGTCTGTTTNTTGTAAAGGTATACACAGATCTCAAGAT	753
QY	214	AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr	233
Db	754	GCAAGGATTTTGGCATAAAAGGACTATAGCTTCAAAAGTGCTCTCTGGTGTGTGTAC	813
QY	234	PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn	253
Db	814	TTCTTTATGAATGGAACCTATGGACTTGCTTTTGGTATGGAACCTCCTTGATTTCTTAAT	873

QY	254	GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer	273								
Db	874	GGAGAACCTGGATATACCATCGGACTGTCTTGTGTTCTTTCTTAGTCTAATCCATAG	933								
QY	274	SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla	293								
Db	934	AGTTATTGCAATGGAGCAGCAGTCCCTCACTTTGAAACCTTCGCAATAGCCGAGAGCT	993								
QY	294	AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla	313								
Db	994	GCCTTTTCATATTTCCAGGTTATTGATAAGAAACCCAGTATAGATAAATTTCCACAGCT	1053								
QY	314	GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr	333								
Db	1054	GGATATAAACCTGAATCCATAGAGGAACCTGTGGAATTTAAATAATGTTCTTTCAATTAT	1113								
QY	334	ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu	353								
Db	1114	CCATCAAGACCATCTATCAAGATTCTGAAAGGTCTGAATCTCAGAAATTAAGTCTGGAG	1173								
QY	354	ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln	373								
Db	1174	ACAGTCGCTTGGTGGTCTCAATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAG	1233								
QY	374	ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu	393								
Db	1234	AGGTTATATGATCCGATGATGGCTTTATCATGGTGGATGAGAATGACATCAGAGCTTTA	1293								
QY	394	AnnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly	413								
Db	1294	AATGTCGGCATTTATCGAGACCATATTTGGAGTGGTTAGTCAAGAGCCTGTTTGTTCGGG	1353								
QY	414	ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu	433								
Db	1354	ACACCATCAGTAAACATATCAAGTATGACGAGATGATGCTACTGATGACGAGATGGAG	1413								
QY	434	ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn	453								
Db	1414	AGAGCAGCAAGGGAAGCAAAATGCTATGATTTTATCATGGAGTTTCTTAATAAATTAAT	1473								
QY	454	ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle	473								
Db	1474	ACATTTGGTAGGGGAAAGAGAGCTCAAAATGAGTGGAGGCGAGAAAACAGAGGATTCGAA	1533								
QY	474	AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu	493								
Db	1534	GTCGTGCTTGTAGTTGCAAAACCCAGAGATTTGATTTAGATGAGGCTAGCTGCGCTG	1593								
QY	494	AspSerGluSerLysSerAlaValGlnAlaLeuGluLys	507								
Db	1594	GATTCAAGAAAGCAAGTCAGCTGTTCAAGCTGCACCTGGAGAAG	1635								

RESULT 5	AX478104	3699 bp	DNA	linear	PAT 12-AUG-2002
LOCUS	Sequence 34 from Patent WO0240541.				
DEFINITION	AX478104				
ACCESSION	AX478104.1	GI:22217064			
VERSION					
KEYWORDS	human.				
SOURCE	Homo sapiens				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	1				
	Tang, Y.T., Yue, H., Nguyen, D.B., Hafalia, A.J., Elliott, V.S., Lu, Y.,				
	Walia, N.K., Yao, M.G., Baughn, M.R., Gandhi, A.R., Ding, L.,				
	Sanjanwala, M., Ramkumar, J., Arvizu, C., Gietzen, K.J., Lal, P.G.,				
	Azimzai, Y., Khan, F.A., Thangavelu, K., Thornton, M., Lu, D.A.,				
	Tribouley, C.M., Warren, B.A., Ison, C.H., Das, D., Raumann, B.E.,				
	Policky, J.L. and Kearney, L.				
	Transporters and ion channels				
TITLE	Patent: WO 0240541-A 34 23-MAY-2002;				
JOURNAL	Incyte Genomics, Inc. (US)				

FEATURES

Location/Qualifiers

source
1. .3699
/organism="Homo sapiens"
/db_xref="taxon:9606"
/notes="Incyte ID No: 7472030CB1"

BASE COUNT 1116 a 707 c 860 g 1016 t
ORIGIN

Alignment Scores:

Pred. No.: 9, 79e-182 Length: 3699
Score: 2460.50 Matches: 495
Percent Similarity: 95.40% Conservatives: 3
Best Local Similarity: 94.83% Mismatches: 9
Query Match: 94.93% Indels: 15
DB: 6 Gaps: 2

US-09-873-409-8 (1-514) x AX478104 (1-3699)

Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 145 ATGATCTGGGTATACACATCACTGTTCAATGGAGCCTGCCTTCTTTAAATGCCACTG 204
Qy 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
Db 205 TGTATAGGAGAAATGAGTGAATACCTTATTAGTGGATGCTAGTCCACACTAACACA 264
Qy 41 SerPhePheArg-----LeuThr 46
Db 265 AATTATCAGAACTGACTCTCAAGAGAGCTGAATGAAGATATGACTCTGTTGACC 324
Qy 47 LeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeu 66
Db 325 CTGATTATGTTGAATAGGTGCTGCTGCTTGTATTTGGTTACATACAGATTCTCTTG 384
Qy 67 TrpIleIleThrAlaAlaArgGlnThrIleArgIleArgIleGlnPhePheHisSerVal 86
Db 385 TGGATTATACTGCAGCAGCACACAGAGAGTTCGAAACAGTTTTCATTTCAGTT 444
Qy 87 LeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGluLeuAsnThrArgMet 106
Db 445 TTGGCAGGAGACATCGGCTGGTGTGATGCTGACATCGGTGAACCTTACACTCGGATG 504
Qy 107 Thr---AspIleAspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGln 125
Db 505 ACAGATGACATTGACAAATCAGTGTGATGAGTATGAGATAAGATTGCTGCTGTTCAA 564
Qy 126 AsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLysGlyTyrLysLeuThr 145
Db 565 AACATGCTACTTTTTCGATTGGCTGGCAGTTGGTTGGTGAAGGGCTGGAACCTCACC 624
Qy 146 LeuValThrLeuSerThrSerProIleMetAlaSerAlaAlaCysSerArgMet 165
Db 625 CTAGTGACTCTATCCAGCTCTCTCTTATATGCTTCAGCGGCAGCATGTTCTAGGATG 684
Qy 166 ValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGlu 185
Db 585 GTCACTCATTCAGCAGTAAGAAATTAAGTCCCTATTCCAAAGCTGGGCTGGCGAGAA 744
Qy 186 GluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGln 205
Db 745 GAAGTCTTGTTCATCAATCCGAACAGTCATAGCCTTTAGGGCCCGAGGAGAAAGACTTCAA 804
Qy 206 ArgTyrThrGlnAnLeuLeuAspAlaLysPheGlyIleLysArgThrIleAlaSer 225
Db 805 AGGTATACACAGAAATCTCAAGATGCAAGGATTTTGGCATTAAGAAAGACTATAGCTTCA 864
Qy 226 LysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTyr 245
Db 865 AAAGTGTCTTGTGCTGTGCTGCTCTTCTTATGATGAAGAACTATGACTGCTTTTGG 924
Qy 246 TyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAla 265
Db 925 TATGGAACTCTCTTGTATCTTAAATGGAGAACTGGATATACATCGGAGCTGTTCTTGCT 984

Qy 266 ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGlu 285
Db 985 GTTTCCTTTAGTGTAAATCCATAGTAGTTATTGTCATTGGAGCAGCAGTCCCTCACTTGA 1044
Qy 286 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPro 305
Db 1045 ACCTTCGAATAGCCGAGAGCTGCTTTTCATATTTTCCAGGTATTATGATAAGAACC 1104
Qy 306 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGlyThrValGlu 325
Db 1105 AGTATAGGTAACTTTTCCAGCTGGATATAAACCTGAATCCATAGAGAACTGTGGAA 1164
Qy 326 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu 345
Db 1165 TTTAAAAATGTTTCTTTCAATATCCATCAAGGCCATCTATCAAGATTCTGAAGGCTCTG 1224
Qy 346 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 365
Db 1225 AATCTCGGAATTAAGTCTCGAGAGACAGTCGCTTGGTCGGTCTCAATGGCAGTGGAA 1284
Qy 366 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 385
Db 1285 AGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCGAGATGCGCTTTATCATGG 1344
Qy 386 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal 405
Db 1345 GATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATGGAGTGT 1404
Qy 406 SerGlnGluProValLeuPheGlyThrIleSerAsnAsnIleLysTyrGlyArgAsp 425
Db 1405 AGTCAAGAGCCTGTTTGTTCGGGACCATCAGTAACAATATCAAGTATGACGAGAT 1464
Qy 426 AspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIle 445
Db 1465 GATGTCACTGATGAAGATGGAGAGCAGCAGGAGGAAAGCAATCGGTATGATTATC 1524
Qy 446 MetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGly 465
Db 1525 ATGAGATTTCTTAATAATTAATACATTTGTTAGGGGAAAAGAGCTCAATGAGTGA 1584
Qy 466 GlyGlnGlnArgIleAlaAlaArgAlaLeuValArgAsnProLysIleLeuIle 485
Db 1585 GGCAGAAAACAGAGGATCGCAATTCGTCGCTTATGTTAGTTCGAAACCCCAAGATTCTG 1644
Qy 486 LeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaLeu 505
Db 1645 TTAGTAGGAGCTACGCTGCTCCCTGGATTGAGAAAGCAAGTCAGCTGTTCAAGCTG 1704
Qy 506 GluLys 507
Db 1705 GAGAAG 1710

RESULT 6

LOCUS AX339030 3177 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 12 from Patent WO0194400.
ACCESSION AX339030
VERSION AX339030.1 GI:18129122

KEYWORDS

human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 Frank, M.H. and Sayegh, M.H.

A gene encoding a multidrug resistance human p-glycoprotein

homologue on chromosome 7p15-21 and uses thereof

Patent: WO 0194400-A 12 13-DEC-2001;

THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)

FEATURES

Location/Qualifiers
1. .3177
/organism="Homo sapiens"

BASE COUNT 970 a 601 c 735 g 870 t 1 others
ORIGIN /db_xref="taxon:9606"

Alignment Scores:
Pred. No.: 11e-122 Length: 3177
Score: 1699.50 Matches: 343
Percent Similarity: 92.70% Conservative: 0
Best Local Similarity: 92.70% Mismatches: 0
Query Match: 65.57% Indels: 27
DB: 6 Gaps: 1

US-09-873-409-8 (1-514) x AX339030 (1-3177)

```
QY 165 MetValIleSerLeuThrSerLysGluSerAlaTyrSerLysAlaGlyAlaValAla 184
Db 1 ATGGTCATCTCATGACCACTAGGAATTAAAGTGCCTATTCCAAAGCTGGGGCTGGCA 60
QY 185 GluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeu 204
Db 61 GAAGAAGTCTTGTCATCAATCCGAACAGTCATAGCCTTTAGGGCCCGAGGAGAAAGACTT 120
QY 205 Gln----- 205
Db 121 CAAGGTCCTTCCTTTAAATATATAACAGATATGCTTGCTGGTTTATTTCCCCAGTGGCTA 180
QY 206 -----ArgTyrThrGlnAsnLeuLysAspAlaLysAspPhe 217
Db 181 CTAAGTTGTGTTCTGTTNTTGTAAAGTATACACAGAACTCTCAAGATGCAAGGATTTT 240
QY 218 GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsn 237
Db 241 GGCATAAAAGGACTATAGCTTCAAAAGTCTCTCTGGTCTGTGTACTTCTTTATGAAT 300
QY 238 GlyThrTyrGlyLeuAlaPheTyrGlyThrSerLeuIleLeuGlnGlyGluProGly 257
Db 301 GGAACCTATGGACTTGCTTTTGTGTATGGAACCTCTCTGTATCTTAAATGGAGAACCTGGA 360
QY 258 TyrThrIleGlyThrValLeuAlaValPheSerValIleHisSerSerTyrCysIle 277
Db 361 TATACCATCGGACTGTTCTTGCTGTTTCTTTAGTGTATTCATCAGCAGTATTGCATT 420
QY 278 GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 297
Db 421 GGAGCAGCAGTCCCTCACTTTGAAACCTTCGCAATAGCCCCGAGGAGCTGCTTTCATATT 480
QY 298 PheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysPro 317
Db 481 TTCAGGTTATTGATAAGAAACCCAGTATAGATAAATCTTTCCACAGCTGGATATAAACCT 540
QY 318 GluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgPro 337
Db 541 GAATCCATAGAGAACTGTGGAAATTTAAATAATGTTTCTTTCAATATCCATCAAGACCA 600
QY 338 SerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu 357
Db 601 TCTATCAAGATTCTGAAAGTCTGAACTCTCAGAAATTAAGCTGGAGAGACAGTCGCCTTG 660
QY 358 ValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAsp 377
Db 661 GTCGGTCTCAATGGCAGTGGGAAGAGTACCGTAGTCCAGCTTCTTCGAGAGGTTATATGAT 720
QY 378 ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis 397
Db 721 CCGGATGATGCTTTATCATGGTGGATGAGAAATGACATCAGAGCTTTAAATGTGGCGCAT 780
QY 398 TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSer 417
Db 781 TATCGACCATATTTGGAGTGGTTAGTCAAGAGCCTGTTTGTTCGGGACCAACCATCAGT 840
QY 418 AsnAsnIleLysTyrGlyArgAspAspValThrAspGluMetGluArgAlaAlaArg 437
Db 841 AACAAATATCAAGTATGGACGAGATGATGTACTGATGAAGAGATGGAGAGACGAAGG 900
```

```
QY 438 GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly 457
Db 901 GAAGCAAAATCGGTATGATTTTATCATGAGAGTTTCCTAATAAAATTTAATACATTGGTAGG 960
QY 458 GluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu 477
Db 961 GAAAAAGGAGCTCAATAGTAGTGGAGGCGAAGAACAGAGATCGCAATTGCTCGTGCTTA 1020
QY 478 ValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 497
Db 1021 GTTCGAACCCCAAGATTCTGATTTAGATGAGGTACGCTGCGCTGGATTCAAGAAAGC 1080
QY 498 LysSerAlaValGlnAlaLeuGluLys 507
Db 1081 AAGTCAGCTGTTCAAGCTGCACCTGGAGAAG 1110
RESULT 7
LOCUS AX322791 4189 bp DNA linear PAT 07-JAN-2002
DEFINITION Sequence 5 from Patent WO0192877.
ACCESSION AX322791
VERSION AX322791.1 GI:18093768
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Sorrentino,B. and Schuetz,J.
TITLE Method of identifying and/or isolating stem cells
JOURNAL Patent: WO 0192877-A 5 06-DEC-2001;
ST. JUDE CHILDREN'S RESEARCH HOSPITAL (US)
FEATURES
Source 1. 4189
/organism="Mus musculus"
/db_xref="taxon:10090"
BASE COUNT 1204 a 875 c 1028 g 1082 t
ORIGIN
Alignment Scores:
Pred. No.: 4.99e-114 Length: 4189
Score: 1590.00 Matches: 307
Percent Similarity: 76.32% Conservative: 99
Best Local Similarity: 57.71% Mismatches: 94
Query Match: 61.34% Indels: 32
DB: 6 Gaps: 5
US-09-873-409-8 (1-514) x AX322791 (1-4189)
QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 148 ATGATTTCTGGAACTCTCGCTGCTATATCCATGGACATTACTTCCCTCTTGATGCTG 207
QY 21 ValLeuGlyGluMetSerAspAsnLeuIle----- 30
Db 208 GTCTTTGGAAACATGACAGATAGTTTACAAAAGCAGAACCCAGTAGTTCTGCAAGCATT 267
QY 31 -----SerGlyCysLeuValGlnThrAsnThr----- 39
Db 268 ACTAATCAAAAGTGGACCCCAACAGTACTCTGATCATCAGCAACAGCAGTCTGGAGGAGAG 327
QY 40 -----TyrSerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAla 56
Db 328 ATGGCATATACGCCCTAC-----TATTACACCGGATTTGGTGTGGTGTG 372
QY 57 LeuIlePheGlyTyrIleGlnIleSerLeuThrPileIleThrAlaAlaArgGlnThrLys 76
Db 373 CTCATAGTTGCCCTACATCCAGGTTTCACTTTGTGCTGTCAGCTGGACACAGATACAC 432
QY 77 ArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyThrPheAspSer 96
Db 433 AAGATTAGGCAGAAAGTTTTTCCATGCTATAATGAATCAGGAGATAGCTGCTGTTGATGTG 492
```



```
/codon_start=1
/protein_id="CAC37723.1"
/db_xref="GI:13921233"
/translation="MDPEGRKGSAEKFNWFKGKSKKKBKKEKPTVSTFAMFRYSN
WLDRLVLMVTWAAIHGAALPMLVFNMTDVSFANAGISRNKTPFVLINESITNT
QHFNLHEEMTYYAYISGIGAGVLAAYIQVSWFCAAAGRIILKIRKQFPFAMRQ
ELGWFDPDVBGLNRLTDDVSKINIGIDKGMFQSIATFTFGFIVGFRTGWLKL
VILASPVGLSAAIAWAKLTISPTDKELLAYAKAGAAVEVLAARTVIAFGOKEL
ERYKNLEBAKGIGIKKAITANISIGMAFLLIYASYALAFWGTGTSLSSEYTIQVL
TVFFSLVIGAFSGASPIEAPANARGAAEIKLIDNPKSIDSTVSKSGHKPDNKG
NLEFKNVSYFSRKEVKLKGKLNKVSQGTVALVGNSSCGKSTVQMLQRLYDPTD
GMVCDIGQIRITNVRHLREITGVWSQEPVLFATTAENIRYGRNVMTMDEIKAVKE
ANAYDFIMKPNKFDLGERGAOLSGQOKQIATARLVRNPKILLDEATSDLTE
SEAVQVADKARKGKTTIVIAHRLSTVENVADVIAGFDGVIKGNHDELWKRGIV
FLKVTMOTRGNIELEENATGESKSDALEMSPKDSGSLIKRSTRSIIHAPQOQDR
KLGTEDLNENPVPSFRWKLKLNSTENFVFGFICAILNGGLQAPRSIIIFSRIGI
FRDEDEPTKQNTNMFVFLVGLIISITFFLQGTGKGAIEILTKRLRYMVRFSM
LRQVSWDFDPKNTGALVRLANDAAOVKAIGSLRAVITQNIANLGTGIIISLYG
WOLTLALLAIVPILTAGVVMKMLSGQALKDKLEGAKIATAEIENFRVWSLTR
EOKPEYMAQSLQVPYRNSLRKAHIFGVFSFIQAMWVPSYAGCFRFGALVNFPMN
FQDVLVFSALVFGMAVQVSPADYAKAKVSAHVMIIEKSPLDISYPHGLKP
NLEGNVTNEVFNTPDIPVJLQGLSLEVKKQTLALVGSSSCGKSTVQVOLLERF
YPLGASVLIDKEIKHLNQLRAHLGIVSQEPILFDCSIAENAYGNSVWVSHHEE
IVQAACEANIHFIETLPKYNTRVGDKGTLGGQOKQIATARLVRNPKILLDEA
TSALDTPSEKVVQEALDKAREGRTCVIAHRLSTIQNADLIVVFQNGKVKERGTQOL
LAQKGIYFSWVQAGAKR"
BASE COUNT 1296 a 833 c 1009 g 1141 t
ORIGIN

Alignment Scores:
Pred. No.: 1-25e-113 Length: 4279
Score: 1595.00 Matches: 304
Percent Similarity: 75.80% Conservative: 97
Best Local Similarity: 57.47% Mismatches: 104
Query Match: 61.15% Indels: 24
DB: 6 Gaps: 4

US-09-873-409-8 (1-514) x AX105082 (1-4279)

QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
DB 167 ATGTGGTGGGCAATGGCTGCATCATCGAGCTGCATCCCTCTCATGATGCTG 226
QY 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu----- 34
DB 227 GTTTTGGAAACATGACAGATAGCTTTGCAAAATGCAGGAATTTCAAGAAACAAACTTTT 286
QY 35 -----ValGlnThrAsnThrThrThrSerPheArg----- 44
DB 287 CCAGTTATATTAATGAAGATATTACGAACAATACACAACTTTCATACCATCTGGAG 346
QY 45 -----LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe 59
DB 347 GAGGAATGACCACTGATGCTATATTACAGTGGGATCGGTGCTGCTGGTGGCT 406
QY 60 GlyTyrIleGlnIleSerLeuTyrIleThrAlaAlaArgGlnThrLysArgIleArg 79
DB 407 GCTTTACATCAGGTTCATCTGCTGCTGGCAGGAGAGACAGATACCAAAATAGA 466
QY 80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99
DB 467 AAACAATTTTTCATGCTATCATGCGACAGAGATGGCTGGTTGACGTGATGAGCTT 526
QY 100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
DB 527 GCGGAGCTTAACACCGGCTCACAGACGATGCTCCAAAATCAATGAAGAAATTGGCGAC 586
QY 119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
DB 587 AAAATTGGAATGTTCTTTCATCAATAGCAACATTTTTCACCGGTTTATAGTGGGGTTT 646
QY 139 ValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
DB 1721 GCAGTGGTTGAGTGGCCCTGGATAAG 1747
```

```
DB 647 ACACGTGGTGGGAAGTAACCCCTTGTGATTTTGGCCATCAGCCCTGTTGACATTTCA 706
QY 159 AlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
DB 707 GCGGCCATCTGGGCAAGATATCTATCTTCTTACTGTATAAGAACTCTTGGCCTATGCA 766
QY 179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198
DB 767 AAAGCTGGAGCAGTAGCTGAGAAGCTTTAGCAGCANTCAGAACTGTGATTGCTTGA 826
QY 199 AlaGlnGluLeuGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 218
DB 827 GGCAAAAGCAAGAACTTCAAAGGTACAAACAAAAATTTAGAAGAACTAAAGAAATTGGG 886
QY 219 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGly 238
DB 887 ATAAGAAGAACTATACGCCCAACATTTCTATGGTGGCCGCTTTCTTATGTACTATGCA 946
QY 239 ThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 258
DB 947 TCATATGCTCTGGCTTCTGGTATGGACCTCTTGGTCTCTCCAGTGAA-----TAT 1000
QY 259 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 278
DB 1001 ACTATTGGACAACTACTACTGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1060
QY 279 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 298
DB 1061 CAGGCATCCCAAGCAATTTGAAGCAATTTGCAAAACGCAAGAGGACGACTTATGAAATCTTC 1120
QY 299 GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 318
DB 1121 AGATATTAATGACATTAACCAAGCAATTTGACGCTATTCAGAGTGCGACATAAACAGAT 1180
QY 319 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 338
DB 1181 AATATTAGGGAATTTGGAATTTCAAAATTTTCACTTCACTTACCTTCTCAGAAAGAA 1240
QY 339 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 358
DB 1241 GTTAAGATCTTAAGGGTCTCAACCTGAAGTTTCAGAGTGGGCGACAGCTGGCGCTGGTT 1300
QY 359 GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro 378
DB 1301 GGGAAACAGTGGCTGCGGGAAGAGCAGCAGCTGCAGCTGATGCAGAGGCTCTATGACCCC 1360
QY 379 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 398
DB 1361 ACAGATGGCATGCTGCTGATTTGATGGACAGGACATTTAGGACCAATAATGTAAGGCATCT 1420
QY 399 ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn 418
DB 1421 CGGAAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1480
QY 419 AsnIleLysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGlu 438
DB 1481 AACATTCGCTATGGCCGCAAAATGTACCATGGATGAGATTGAGAAAGCTGTATAGGAA 1540
QY 439 AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 458
DB 1541 GCCAATGCTATGATTTTATCATGAAACTACCTATAAATTTGACACTCTGCTGTTGAGAG 1600
QY 459 LysGlyAlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal 478
DB 1601 AGAGGGGCCAGCTGAGTGGTGGACAAACAGAAATGCCCAATGCTCGGGCCCTGGTT 1660
QY 479 ArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 498
DB 1661 CGCAACCCCAAGATTTCTTCTGCGATGAGGACAGCTCAGCTCTGACACTGGAAGTGAA 1720
QY 499 SerAlaValGlnAlaLeuGluLys 507
DB 1721 GCAGTGGTTGAGTGGCCCTGGATAAG 1747
```


RESULT 10	AX105078	4279 bp	DNA	linear	PAT 30-APR-2001
LOCUS	AX105078	22 from Patent	W00123540.		
DEFINITION	AX105078				
ACCESSION	AX105078				
VERSION	AX105078.1	GI:13921228			
KEYWORDS					
SOURCE	dog.				
ORGANISM	Canis familiaris				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.				
REFERENCE	1 (bases 1 to 4279)				
AUTHORS	Stocker,P.J., Steinel-Crespi,D.T., Crespi,C.L., Reif,T.C. and Patten,C.J.				
TITLE	P-glycoproteins and uses thereof				
JOURNAL	Patent: WO 0123540-A 22 05-APR-2001;				
	GENTEST CORPORATION (US)				
FEATURES	Location/Qualifiers				
source	1..4279				
	/organism="Canis familiaris"				
	/db_xref="taxon:9615"				
CDS	17..3862				
	/note="unnamed protein product"				
	/codon_start=1				
	/protein_id="CAC317721"				
	/db_xref="GI:13921229"				
	/translation="MDPEGKGSAEKNFWMGKKSKNKKKKKPTVSTFAMFRVSN				
	WLDRLVLMVGTMAAITHGAALPLMLVLFGNMTDSFNAGISRNKTFFVIINESITNT				
	QHPINLHEEMTTVAYYSIGAGVLVAAYIQVSWFCLAGROIILKIRQFPAHMRQ				
	EIGWFDVHDVGLNTRTDDVSKINIGDKIGMPFOSIATFTFTGLVGRGKWLTL				
	VILAISSPLGSLNAIKWILSKSTFDKELLAYAKAGAAVEELAAIRTVIAFGGKKKL				
	ERYNNLEAKGIGIKAITANISIGAFLLIYASVALAFWYGTSLVLSSEYSIGQVL				
	TFVSVLIGAFSGIQASPSIEAFANARGAAAEYIFKIIDNPKSIDSKSGHKPDNLK				
	NLFSPGNHFYSPRKEVYKILKLNKLVQSGQTVALVNSGCGKSTTVQLMQRDLYG				
	GMVICDQDIRTINVRHLREITGVSQEPLVFATTIAENIRYGRENVTMDEATSKVE				
	ANAYDFTMKLPNKFDTLVUGRGAQLSGQKORIAARALVRNPKILLDEATSAKTE				
	SEAVVOALDKASKGRFTTVIAHRLSTVRNADVLAGDDGVIYVEKGNHDELMEKGYI				
	PKLVMTQTRNETELENAATESKSEDALESPKDSGSSLLIKRSTERSIHAQGGDR				
	KLGTKEDLNENFVFWRLIKLNSTWPIYVVGIFCAIINGGLQPAFSIIFSRIGI				
	LTRQDSVPDTQRNSMFSVFLVLGIIISFTFFLQGTFGKAGEIILTKLRYVMVFSRM				
	FRRQDSVPDTQRNSMFSVFLVLGIIISFTFFLQGTFGKAGEIILTKLRYVMVFSRM				
	WQTLALLAIVPIIAIAGVEMKLSQALKDKKELEGAKIATAEINERTVVLSSLY				
	EOKFPEYWAQSLQVPVYRNSLRKAKHIFGVFSFIQAMMYFSYAGCFRFGALVAFNEFN				
	PDQVLLVPSAIVFGMAAGVQSSPADPYAKAKVSAAHVIMIEKSPIDSYSPHGLKP				
	NLTLEGNTFVNEVFNTRDPIPVQLGSLIEVSKGQTLALVSGSGCKSTVQQLLERP				
	YDPLAGSLVDIGKEIKHLNVQMLRAHIGIYVQEPILFDCAENAIARALVRPHILLDEA				
	IMQAKEANTHHFIETLPKRYNTRVGDKGTQLSGQKORIAARALVRPHILLDEA				
	TSALDSEKVVQELDKAREGRTCVIAHRLSTIQNADLIIVFQNGKVKHEGHTHOOL				
	LAQGIYFMSVMSVQAGAKR"				
BASE COUNT	1295 a	833 c	1008 g	1143 t	
ORIGIN					
Alignment Scores:					
Pred. No.:	2,56e-113	Length:	4279		
Score:	1581.00	Matches:	303		
Percent Similarity:	75.80%	Conservative:	98		
Best Local Similarity:	57.28%	Mismatches:	104		
Query Match:	61.00%	Indels:	24		
DB:	6	Gaps:	4		
US-09-873-409-8 (1-514) x AX105078 (1-4279)					
Qy	1	MettlleLeuGlylleLeuAlaSerLeuValAanGlyAlaCysLeuProLeuMetProLeu	20		
Db	167	ATGTTGGGGCAATGGCTGCCATCATCATCATGAGTGCACTCCCTCATGATGCTG	226		
Qy	21	ValLeuGlyGluMetSerAsnLeulleSerGlyCysLeu	34		
Db	227	GTTTGTGGAAACACAGATAGCTTTTGCMAATGCAGGAAATTCAGAAACAAACTTTT	286		
Qy	35	-----ValGlnThrAsnThyTrSerPhePheArg	44		

QY 279 AlaAlaValProHisPheGluThrPheAlaIleAlaAArgGlyAlaAlaPheHisIlePhe 298
DQ 1061 CAGGCATCCCAAGCATTTGAAGCATTTGCAAAAGCGAAGAGGAGCAGCTTATGAAATCTTC 1120
QY 299 GlnValIleAspLysLysProSerIleAspAenPheSerThrAlaGlyTyrLysProGlu 318
DQ 1121 AGATTAATTGCAATAAACAGCATTTGAGCTATTCGAAGGTGAGCATTAACAGAT 1180
QY 319 SerIleGluGlyThrValGluPheLysAenValSerPheAsnTyrProSerArgProSer 338
DQ 1181 AATATTAAAGGAAATTTGGAATTCAAATGTTCACTTCAGTTACCTCTCGAAGAA 1240
QY 339 IleLysIleLeuLysGlyLeuAenLeuArgIleLysSerGlyGluThrValAlaLeuVal 358
DQ 1241 GTTAAGATCTTAAAGGCTCTCAACTGAAGGTTTCAAGTGGGAGGAGGAGTGGT 1300
QY 359 GlyLeuAsnGlySerGlyLysSerThrValGlnLeuLeuGlnArgLeuTyrAspPro 378
DQ 1301 GGAACAGTGGCTGCGGGAAGAGCACGCGTGCAGCTGATGAGAGGCTCTATGACCCC 1360
QY 379 AspAspGlyPheIleMetValAspGluAenAspIleArgAlaLeuAenValArgHisTyr 398
DQ 1361 ACAGATGGCATGCTGTATTGATGACAGGACATTTAGGACCATTAATGTAAGCATCTT 1420
QY 399 ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAen 418
DQ 1421 CGGGAATTAAGTGGTGTGAGTCAAGGAGCTGTGTTGTTGCCACGAGGCTCTATGAG 1480
QY 419 AsnIleLysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGlu 438
DQ 1481 ACATTCGCTATGCGCGCAATATGACATGATGAGATGAGATGAGAAAGCTGTAAAGAA 1540
QY 439 AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 458
DQ 1541 GCCATGCTATGATTTTATCATGAACCTACCTAATAATTTGACACTCTGTTGAGAG 1600
QY 459 LysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal 478
DQ 1601 AGAGGGCCGAGTGTGAGTGGTGGACAGAAACAGAGAAATCGCCATTCGCGGCCCTG 1660
QY 479 ArgAsnProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 498
DQ 1661 CGCAACCCCAAGATTTCTTCGTGGATGAGGCAACGTCAGCTCTGGACACTGAAAGTAA 1720
QY 499 SerAlaValGlnAlaLeuGluLys 507
DQ 1721 GCAGTGGTTTCAGTGGCCCTGGATAAG 1747

RESULT 12
AX108656 LOCUS 4195 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 3 from Patent WO0123565.
ACCESSION AX108656
VERSION AX108656.1 GI:13923888
KEYWORDS crab-eating macaque.
SOURCE Macaca fascicularis
ORGANISM Macaca fascicularis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
Cercopitheciinae; Macaca.
1 (bases 1 to 4195)
REFERENCE Stocker,P.J., Steimel-Crespi,D.T. and Crespi,C.L.
AUTHORS P-glycoproteins from macaca fascicularis and uses thereof
TITLE Patent: WO 0123565-A 3 05-APR-2001;
JOURNAL GENEST CORPORATION (US)
FEATURES
source
1. 4195
/organism="Macaca fascicularis"
/db_xref="taxon:9541"
CDS
100..3951
/note="unnamed protein product"
/codon_start=1

/protein_id="CAC37765.1"
/db_xref="GI:13923888"
/translation="MDLEGDRNGAEKKNFFKLNKSKKDKKRRKPTVSVFSMPFRYSN
WLDKLYMVVGTAAAIHHGAGPLGMLVFGDMTDTFPAAGNGLDGLGALLFNNTSNIT
DTPVKNLEDMTRYAYYSIGAGLVAAVQVSWFACLAAGROGHIKRGQFFHAIIR
QEIWFVDVHGLNRLTDVSKINEGIDKIGMFOSMATFFFTGTVTFVFGWKLK
LVAIAPVLGLSAAWAKILSSFTDKELLAYAKAGAAEVLAIATVTFVFGWKLK
LERNKLEAKRIGIKAITANISIGAAFLIYASALAFWYGTTLVLSKEYSGQV
LTVFVSLVIGAFSQQASPSIEAFANARGAAFEIPKIDNKPISDYSKSGHKPNIK
GNLFRNVHFSYPSRKEVKILKLNKLVQSGQTVLVNCGSGKSTTVQMLQRLDPT
EGMVSDQDRTINVRFLREIIGVSOEPLVFATTAENIRYGREVDVMDIEKAVK
EANAVDTIMKLPQKFDLVGRGAQLSGGOKORAIARALVRNPKILLLDEATSLDT
ESEAVQVALDKARKGRTIIVIAHRLSTVRNADVIAGFDGVIKEGHNDELMEKGI
YFKLVTMTAGNEIELENADESEIDTEMSSHDSSLRKRSRSTSVRSVSGQD
RKLSTKEALDESIPVSWFRIMKLNLTWPFVVGVCALINGGLOPAPAFVFSKIIG
IFTRNDAAETKRONSLFLLVLGIVSFITFLQGTGFGKAGELTKRLRYMVPFR
MLRODVSFDDPKNTTGATTLANDAAOVKAIGSLAIITQNLANLGTGIIISLIY
QMOLTLALLAIVPIIATAGVVMKLSQALKDKELGAGKIAETENPRVTVSLT
QEOKPEHYDQSLQVPRNSLRKAHIFGITEFTQMMYFSGAGCFRFGAIVLHSLM
SFEDLLVFSAVFGMAVGVSSFPADYAKAKVSAHIIMIEKTPILDSYSTGLK
PNTLEGNVTNEVFNFTRLDI PVQLGSLVKKGTALVLSGSGCGKSTVQVLLER
FYDLAGKVLDDKEIKQLNWLRAHLGIVSQEPIFLDCSISENIATGQDRSVSQE
BIVRAAKEANTHAFIESPNKYSTRVGDKGTOLSGGOKORAIARALVRPHILLDE
ATSLDTESEKVOEALDKAREGRTICIVIAHRLSTIQNLADLIIVFQNGRVEHGTQ
LLAQKGIYFSMVSVQAGAKRQ"

BASE COUNT 1230 a 801 c 1039 g 1125 t
ORIGIN

Alignment Scores:
Pred. No.: 2,74e-113 Length: 4195
Score: 1580.50 Matches: 306
Percent Similarity: 75.85% Conservative: 96
Best Local Similarity: 57.74% Mismatches: 103
Query Match: 60.98% Indels: 25
DB: 6 Gaps: 4

US-09-873-409-8 (1-514) x AX108656 (1-4195)

QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
DQ 250 ATGTTGGTGGGAACCTTTGGCTGCCATCATCCATGGAGCTGGACTTCCTCTCATGATGCTG 309
QY 21 ValLeuGlyGluMetSerAspAsnLeuIleSer-----GlyCys 33
DQ 310 GTGTTGGAGACATCGGATACCTTTTGCAATGTCAGGAAATTTAGGAGATTTAGGAGCT 369
QY 34 LeuValGlnThrAsnThrTyrSer-----41
DQ 370 CTGTTGTTTAAACACACTAATAGCAGTAAATACATGATACAGTGCCTGATGAATCTG 429
QY 42 -----PhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIle 58
DQ 430 GAGGAAGATATGACAGGATGATGCCTATTATTACAGTGGAAATGGTCTGGGTGCTGTT 489
QY 59 PheGlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIle 78
DQ 490 GCTGCTTACATTCAGGTTTCATTTGGTGCCTGCGAGCTGGAAGACAAATACACAAAT 549
QY 79 ArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAsp 98
DQ 550 AGAAACACGTTTTTTCATGCTATAATGCGACAGGATAGGCTGGTTTATGATGACAGCT 609
QY 99 IleGlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGly 117
DQ 610 GTTGGGAGCTTAAACCCCGGCTTACAGATGATGCTCCAGATTAATAGGAATGGT 669
QY 118 AspLysIleAlaLeuLeuPheGlnAenMetSerThrPheSerIleGlyLeuAlaValGly 137
DQ 670 GACAAATTTGGAATGTTCTTTCAGTCAATGCGAACACATTTTTCACCTGGTTTATAGTAGGA 729
QY 138 LeuValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAla 157
DQ 730 TTTACACGTGGTGGAAAGCTAACCCCTTGTGATTTTGGCCATCAGCTCTGTTCTTGCAGCTG 789


```
Db 329 GTTTTGGAAACATGACAGATAGCTTTGCAAAATGCGAGGAATTTCAAGAAACAAAACTTTT 388
Qy 35 -----ValGlnThrAsnThrTyrSerPhePheArg ----- 44
Db 389 CCAGTTAATAATTAAGAAAGTATTAGCAACAATATACAACTTTTCATCAACCATCTGGAG 448
Qy 45 -----LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe 59
Db 449 GAGGAAATGACCACGATATGCTATTATTATACAGTGGGATCGGTCTGCGTCTGGTGGCT 508
Qy 60 GlyTyrIleGlnIleSerLeuThrIleIleThrAlaAlaArgGlnThrIleArgIleArg 79
Db 509 GCTTACATCCAGTTTCATCTCGTCCCTGGCAGCAGGAACAGATACATCAAAATTAGA 568
Qy 80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIle 99
Db 569 AAACAATTTTTCATGCTATCATGCGACAGAGGATGCGTGTTCAGCGTCATGACGTT 628
Qy 100 GlyGluLeuAsnThrArgMetThr---AspIleAspIleSerAspGlyIleGlyAsp 118
Db 629 GGGGAGCTTTAAACACCGCTCACAGAGATGCTCCAAAATCAATGAAGAAATGGCGAC 688
Qy 119 LysIleAlaLeuLeuPheGlnIleAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
Db 689 AAAATTGGAATGCTTTCAATCAATAGCAACATTTTTCACCGGTTTATATAGTGGGGTTT 748
Qy 139 ValLysGlyTyrPheLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
Db 749 ACACGGTGTGGAAGTAAACCTTGATTTTGGCCATCAGACCTGCTTCTTGGCATTTCA 808
Qy 159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
Db 809 GCGGCCATCTGGGCAAGATACCTCTCATTTACTGATAAAGAACTCTTGGCCTATGCA 868
Qy 179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198
Db 869 AAAGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGAACTGTGATGCTTTGGA 928
Qy 199 AlaGlnGluLeuGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 218
Db 929 GGAAAGAAAGAAAGAACTTGAAGAGTCAACAAAATTTAGAAAGCTAAAGAAATTTGGG 988
Qy 219 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGly 238
Db 989 ATAAAGAAAGCTATCACGGCAACATTTCTATTGTTGCGCTTCTTATTGATCATGCA 1048
Qy 239 ThrTyrGlyLeuAlaPheTyrPyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 258
Db 1049 TCATATGCTCTGGCTTCTGTTATGGGACCTCCTTGTGCTCTCCAGTGAA-----TAT 1102
Qy 259 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 278
Db 1103 TCTATTGGGCAAGTACTCTACTGCTCTTCTTCTGTTATTAATGCGGCTTTTATGATGGA 1162
Qy 279 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaPheHisIlePhe 298
Db 1163 CAGGCATCCCAAGCATTGAGCAATTGCAACGCAAGAGAGCAGCTTATGAAATCTTC 1222
Qy 299 GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 318
Db 1223 AAGATAATTGACAATAAACCACGATTTGACAGCTATTGGAAGAGTGACATAAACAGAT 1282
Qy 319 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 338
Db 1283 AATATTAAAGGGAATTTGGAATTTCAAAATATGTTCACTTCAGTTACCTTCTCGGAAAGAA 1342
Qy 339 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 358
Db 1343 GTTAAGATCTTAAGGGTCTCAACCTGAGGTTGAGAGTGGGAGAGCAGTGGCGCTGGTT 1402
Qy 359 GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro 378
```

```
Db 1403 GGGAAACAGTGGCTGCGGGGAAGAGCAGCAGCGTGCAGCTGATGCAGAGGCTCTATGACCCC 1462
Qy 379 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 398
Db 1463 ACAGATGGCATGCTGTATTGATGGACAGGACATTAGGACCAATAATGTAAGGCATCTT 1522
Qy 399 ArgAspHisIleGlyValValSerGlnProValLeuPheGlyThrThrIleSerAsn 418
Db 1523 CGGGAATTTACTGGTGTGTGAGTCAGGAGCCTGTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1582
Qy 419 AsnIleLysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGlu 438
Db 1583 AACATTGCTATGCGCGCAAAATGTCACCATGATGAGATGAGAAAGCTGTTAGGAA 1642
Qy 439 AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 458
Db 1643 GCCAATGCTATGATTTTATCATGAAACTACCTAATAAATTTGACACTCTGTTGGTGGAG 1702
Qy 459 LysGlyAlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal 478
Db 1703 AGAGGGGCGCAGCTGAGTGGTGGACAGAAACAGAGAATCGCCATTCCTCGGGCCCTGGTT 1762
Qy 479 ArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 498
Db 1763 CGCAACCCCAAGATTTCTCTGCTGATGAGCAACGTCAGCTCTGGACACTGGAAGTGAA 1822
Qy 499 SerAlaValGlnAlaLeuGluLys 507
Db 1823 GCAGTGGTTCAGTGGCCCTGGATAG 1849
RESULT 14
AX322787
LOCUS AX322787 3860 bp DNA linear PAT 07-JAN-2002
DEFINITION Sequence 1 from Patent WO0192877.
ACCESSION AX322787
VERSION AX322787.1 GI:18093766
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Sorrentino, B. and Schuetz, J.
TITLE Method of identifying and/or isolating stem cells
JOURNAL Patent: WO 0192877-A 1 06-DEC-2001;
ST. JUDE CHILDREN'S RESEARCH HOSPITAL (US)
FEATURES
source 1..3860
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 1135 a 746 c 957 g 1022 t
ORIGIN
Alignment Scores:
Pred. No.: 3,23e-113 Length: 3860
Score: 1579.00 Matches: 306
Percent Similarity: 77.04% Conservative: 100
Best Local Similarity: 58.06% Mismatches: 99
Query Match: 60.92% Indels: 22
DB: 6 Gaps: 5
US-09-873-409-8 (1-514) x AX322787 (1-3860)
Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 151 ATGGTGGTGGAACTTTGCTGCCATCCATCCATGCGGCTGGAGCTTCTCTCATGATGCTG 210
Qy 21 ValLeuGlyGluMetSerAsp-----AenLeuIleSer 31
Db 211 GTGTTTGGAGAAATGACAGATATCTTTGCAATGACAGAAATTTAGAGATCTGATGCTCA 270
Qy 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
```


D	b	271	AACATCACAATAAGAGTGATATCAAAAGATACAGGGTTCTTCTCATGATCTGCAGCAAGAC	330
Q	y	44	-----ArgLeuThrLeuTyTrValGIyIleGIyValAlaAlaLeuIlePheGIyTr	61
D	b	331	ATGACCAGATATCCCTATTATTACAGTGAATTGGTGTGGGTCTGGTGTCTTC	390
Q	y	62	IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln	81
D	b	391	ATTCCAGGTTCATTTTGGTCTCGGACGCTGGGAAGACAAATACACAAATATTAGAAAA	450
Q	y	82	PhePheHisSerValLeuAlaGlnAapIleGIyTrpPheAspSerCysAspIleGIyGlu	101
D	b	451	TTTTTTTCATGCTAATATGACAGGAGTAGCTGTGTGATGTCACAGATTGTGGGAG	510
Q	y	102	LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGIyIleGIyAspLysIle	120
D	b	511	CTTAACACCCGACTTACAGATGATGCTCTAGATTAATGAAGGTATTGGTGACAAAATT	570
Q	y	121	AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGIyLeuAlaValGIyLeuVallys	140
D	b	571	GGAATGTCTTTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGGATTTACAC	630
Q	y	141	GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla	160
D	b	631	GGTTGGAAGCTAACCCCTGTGATTTTGGGCCATCAGTCTCTGTGTTTGGAGCTGCT	690
Q	y	161	AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyTrSerLysAla	180
D	b	691	GTCGTGGCAAAGATACATCTCTTCTTACTATGAAGAAGAACTCTTAGCTGTAGCAAA	750
Q	y	181	GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln	200
D	b	751	GGACGAGTAGCTGAAGAGGCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGCAA	810
Q	y	201	GluLysGlnLeuGlnArgTyThrGlnAsnLeuLysAspAlaLysAspPheGIyLys	220
D	b	811	AAGAAGAAGCTTGAAGAAGGTACAAACAAAAATTTAGAGAAGAGCTTAAAGAATTT	870
Q	y	221	ArgThrIleAlaSerLysValSerLeuGlyAlaValTyTrPhePheMetAsnGlyThrTy	240
D	b	871	AAAGCTAATTACAGCCAATATTTCTATAGGTGTGCTTCTCTCTCATCTATGCATCT	930
Q	y	241	GlyLeuAlaPheTrpTyTrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyThrIle	260
D	b	931	GCTCTGGCTTCTGGTATGGCACCACTGGTCTCTCTCAGGGGAA-----TATTCTATT	984
Q	y	261	GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyTrCysIleGIyAlaAla	280
D	b	985	GGACAAGTACTCACTGTATTCTTTTCTGTATTAAATTTGGGGCTTTTTGTGTGGACAG	1044
Q	y	281	ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal	300
D	b	1045	TCTCCAAGCAATTGAAGCATTTGCAAAATGCAAGAGGACGAGCTATTGAANAATCTT	1104
Q	y	301	IleAspLysLysProSerIleAspAsnPheSerThrAlaGIyTyLysProGlySerIle	320
D	b	1105	ATTGATTAATGAACCAAGTATTGACAGGTATTGCAAGAGTGGGCACAAACCAAGATA	1164
Q	y	321	GluGlyThrValGluPheLysAsnValSerPheAsnTyTrProSerArgProSerIleLys	340
D	b	1165	AAGGAAAATTTGGAATTCAGAAATGTTCACCTTCAGTTACCCATCTCGAAAAGAGTTAA	1224
Q	y	341	IleLeuLysGIyLeuAsnLeuArgIleLysSerGIyGluThrValAlaLeuValGIyLeu	360
D	b	1225	ATCTTGAAGGCGCTGAACCTTGAAGGTGCAGAGTGGGCACACGGTGGCCCTGGTGGAA	1284
Q	y	361	AsnGlySerGIyLysSerThrValValGlnLeuLeuGlnArgLeuTyTrAspProAsp	380
D	b	1285	AGTGGCTGTGGGAACAGACACACACNGTCCAGTCAATGCAGAGGCTCTATGACCCCAC	1344
Q	y	381	GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyTrArgAsp	400
D	b	1345	GGGATGGTGCAGTGTGTATGACAGGATATTAGACCACTAAATGTAAAGTTTCTACGG	1404

QY	401	HisleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAseIle	420
Db	1405	ATCATTTGGTGGTGGAGCAGAACTGTATTGTTGCCACACAGATAGCTGGAACCAATT	1464
QY	421	LysTyGlyArgAspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsn	440
Db	1465	CGCTATGGCCGTGAAATGTCACCATGATGAGATTGAGAAAGCTGTCAAGAGAGCCAAT	1524
QY	441	AlaTyAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGly	460
Db	1525	GCCTATGACTTTATCATGAAACTGCTCATAAATTTGCACACCTGGTTGGAGAGAGAGGG	1584
QY	461	AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn	480
Db	1585	GCCCAAGTTGAGTGGTGGCGCAAGACAGAGGATCGCCATTGACGTGCCCTGGTTGCGCAAC	1644
QY	481	ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla	500
Db	1645	CCCAAGATCTCTCTGCTGATGAGGCCACGTTCAGCCTTGGACACAGAAAGCGAGACGTG	1704
QY	501	ValGlnAlaAlaLeuGluLys	507
Db	1705	GTTCAAGTGGCTCTGGATAAG	1725
RESULT	15		
E02326			
LOCUS			
DEFINITION	E02326	4378 bp	RNA linear
ACCESSION	E02326	Multi drug resistance relating gene derived from human normal cells.	PAT 29-SEP-1997
VERSION	E02326.1	GI:2170561	
KEYWORDS	JP 1990100680-A/1.		
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1. (bases 1 to 4378)		
TITLE	Ueda, K. and Komano, T.		
JOURNAL	HUMAN NORMAL CELL-DERIVED MDR RELATED GENE		
COMMENT	Patent: JP 1990100680-A 1 12-APR-1990;		
	SUNTORY LTD		
	OS Homo sapiens		
	PN JP 1990100680-A/1		
	PD 12-APR-1990		
	PF 05-OCT-1988 JP 1988251475		
	PI UEDA KAZUMITSU, KOMANO TORU		
	PC C12N15/12, C12N1/21, C12Q1/68;		
	CC strandedness: Single;		
	CC topology: Linear;		
	CC *source: tissue type=Adrenal gland;		
	CC *source: clone=SM1132;		
	FH Key	Location/Qualifiers	
	FH		
	FT 5'UTR	1. .137	
	FT CDS	138. .3980	
	FT	/gene='Multidrug resistance relating gene' FT	
	FT mat_peptide	138. .3977	
	FT	/gene='Multidrug resistance relating gene' FT	
	FT 3'UTR	/genes='MDRAL'	
		3981. .54378.	
FEATURES	Location/Qualifiers		
source	1. 4378		
	/organism='Homo sapiens'		
	/db_xref='taxon:9606'		
BASE COUNT	1315 a 818 c 1057 g 1188 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	3.77e-113	Length:	4378
Score:	1579.00	Matches:	306
Percent Similarity:	77.04%	Conservative:	100
Best Local Similarity:	58.06%	Mismatches:	99
Query Match:	60.92%	Indels:	22


```
DB: 6 Gaps: 5
US-09-873-409-8 (1-514) x E02326 (1-4378)
Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAenGlyAlaCysLeuProLeuMetProLeu 20
Db 289 ATGGTGGTGGGAACCTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 347
Qy 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
Db 348 GTGTTGGAGAAATGACAGATATCTTGCAAATGCAGGAAATTTAGAAAGATCTGATGTCA 407
Qy 32 GlyCysLeuValGlnThr-----AenThrTyrSerPhePhe----- 43
Db 408 AACATCACTAATAGAAAGTATATCATGATATACAGGCTTCTTCATGAATCTGGAGGAAGAC 467
Qy 44 -----ArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db 468 ATGACCAAGTATGCCCTATTATTACAGTGAATTTGGTGGGGTGTCTGCTGCTTAC 527
Qy 62 IleGlnIleSerLeuThrIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db 528 ATTCAGTTTCATTTGGTGGCTGGCAGCTGGAGACAATACACAAATTAGAAAACAG 587
Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101
Db 588 TTTTTCATGCTATATGCGACAGGAGATAGGCTGTTTGTATGTGCACCATGTTGGGGAG 647
Qy 102 LeuAenThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db 648 CTTAACACCCGACTTACAGATGATGCTCCAAGATTAAATGAAGGAATTTGGTACAAAAT 707
Qy 121 AlaLeuLeuPheGlnAenMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db 708 GGAATGTTCTTTCAGTCATAGGCAACATTTTCACCTGGGTTATAGTAGATTACACGT 767
Qy 141 GlyTyrPheLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db 768 GGTGGGAAGCTAACCTTGTGATTTTGGCCATCAGTCTCTGTTCTTGGACTGTGCAGTGTCT 827
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db 828 GTCTGGGCAAGATACATCTTTCATTTACTGATAAAGAACTCTTACGCTATGCAAAAGCT 887
Qy 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db 888 GGAGCAGTAGCTGAAGAGGCTCTGGCAGCAATTTAGCACTGTGATTCATTTGGAGGACAA 947
Qy 201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
Db 948 AAGAAAGAACTTGAAGAGGTACACAAATAATTTAGAAGAGCTTAAAGAAATTTGGGATAAAG 1007
Qy 221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAenGlyThrTyr 240
Db 1008 AAGAGCTATTACAGCAATATTTCTATAGGTGCTGCTTCTCTGCTGATCTATGCATCTTAT 1067
Qy 241 GlyLeuAlaPheThrTyrGlyThrSerLeuIleLeuAenGlyGluProGlyTyrThrIle 260
Db 1068 GCTCTGGCTTCTGGTATGGGACACCTTGGTCTCTCAGGGGAA-----TATTCATTT 1121
Qy 261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
Db 1122 GGACAAAGTACTCACTGTATTCTTTCTGTATTAAATTTGGGGCTTTTAGTGTGGACAGGCA 1181
Qy 281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
Db 1182 TCTCCAAGCATTTGAAGCATTTGCAAAATGCAAGAGGAGCAGCTTATGAAATCTTTCAAGATA 1241
Qy 301 IleAspLysLysProSerIleAspAenPheSerThrAlaGlyTyrLysProGluSerIle 320
Db 1242 ATTTGATAATAAGCAAGATTTTGACAGCTATTGCAAGAGTGGGCAACCAACCATATAATT 1301
Qy 321 GluGlyThrValGluPheLysAsnValSerPheAenTyrProSerArgProSerIleLys 340
```

```
Db 1302 AAGGGAATTTGGAATTCAGAAATGTTTCACTTCAGTTACCCATTCGAAAAGAAAGTTAAG 1361
Qy 341 IleLeuLysGlyLeuAenLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
Db 1362 ATCTTGAAGGTTCTGAACCTGAAGCTGCAGAGTGGGCAGAGCGTGGCTGTTGGAAAC 1421
Qy 361 AenGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
Db 1422 AGTGGCTGTGGGAAGAGCACACACATCCAGCTGTATGCGAGGCTCTATGACCCACAGAG 1481
Qy 381 GlyPheIleMetValAspGluAenAspIleArgAlaLeuAenValArgHisTyrArgAsp 400
Db 1482 GGGATGGTCACTGAGTGTGGAGCAGGATATTAGGACCATAAATGTAAAGGTTTCTACGGAA 1541
Qy 401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
Db 1542 ATCATTTGGTGTGGTGGTGGAGAACTGTATTGTTTGGCCACCGATAGCTGAAAACATT 1601
Qy 421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
Db 1602 CGCTATGGCCGTGAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 1661
Qy 441 AlaTyrAspPheIleMetGluPheProAenLysPheAenThrLeuValGlyGluLysGly 460
Db 1662 GCCTATGACTTTTATCATGAACTGCCTCATAAATTTGACACCTGTTGGAGAGAGGG 1721
Qy 461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db 1722 GCCCAGTTGAGTGGTGGGAGAGGAGGAGGATCGCATTTGCAGCTGCTGTTGCTAAC 1781
Qy 481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAenSerGluSerLysSerAla 500
Db 1782 CCCAAGATCCTCTGCTGGATGAGGCCAGCTCAGCCTTGGACACAGAAAGCGAAGCAGTG 1841
Qy 501 ValGlnAlaAlaLeuGluLys 507
Db 1842 GTTCAGGTGGCTCTGGATAAG 1862
```

Search completed: March 31, 2003, 02:24:15
Job time : 7000.72 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4 p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 30, 2003, 01:20:01 ; Search time 502.349 Seconds
(without alignments)
2304.231 Million cell updates/sec

Title: US-09-873-409-8

Perfect score: 2592

Sequence: 1 MILGILASLVNGACILPLMLP.....SESKSAVQAALKEKTPRYSF 514

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=xlh
-Q/cn2_1/USPTO_spool/US9873409/runat_27032003_115416_19216/app_query.fasta_1.7544
-DB=N_Geneseq_101002 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US9873409 @CN1_1_4247 @runat_27032003_115416_19216 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -FgaPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_101002.*
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Length	DB ID	Description
1	2460.5	94.9	3699	24	ABK83223 Human transporter
2	1590	61.3	4189	21	AAZ49334 Murine multidrug r
3	1590	61.3	4189	24	ABA94367 Mouse BCRP DNA rel
4	1590	61.3	4313	14	AAQ38950 Mouse multidrug re
5	1585	61.1	4279	22	AAQ38950 Dog P-glycoprotein
6	1581	61.0	4279	22	AAQ38950 Dog P-glycoprotein
7	1581	61.0	4279	22	AAQ38950 Dog P-glycoprotein
8	1580.5	61.0	4195	22	AAQ38950 Cynomologous monke
9	1579	60.9	3860	21	AAZ49332 Human wild-type mu
10	1579	60.9	3860	24	ABA94365 Human BCRP DNA rel
11	1576	60.8	4233	21	AAZ90198 Rat mdr1b2 (multis
12	1576	60.8	4233	22	AAF27498 Rat mdr1b2 multidr
13	1576	60.8	4279	22	AAQ38950 Dog P-glycoprotein
14	1576	60.8	4317	22	AAQ38950 Dog P-glycoprotein
15	1576	60.8	4669	19	AAV32645 Human P glycoprote
16	1576	60.8	4669	24	ABK52041 cDNA encoding huma
17	1572	60.6	4186	22	AAF86127 Cynomologous monke
18	1571.5	60.6	4254	24	ABK63517 Rat sequence diffe
19	1570	60.6	3860	21	AAZ49333 Human G185V mutat
20	1570	60.6	3860	24	ABA94366 Human BCRP DNA rel
21	1570	60.6	4349	22	AAH57442 Human intestine ce
22	1570	60.6	4646	15	AAQ72872 Human multidrug re
23	1570	60.6	4646	21	AAZ94738 Human ATP binding
24	1570	60.6	4646	24	AAZ38994 Human mdr1 gene.
25	1570	60.6	4646	24	ABL68880 Kidney cancer rela
26	1570	60.6	4646	24	ABL68880 Hybrid vector p8F
27	1570	60.6	6505	17	AAT11394 Retroviral M4 mdr-
28	1570	60.6	8630	21	AAZ24041 Retroviral vector
29	1570	60.6	8630	21	AAZ24042 Mutated human P-g1
30	1568.5	60.5	4264	19	AAV66533 Sequence of human
31	1568.5	60.5	4264	19	AAV66533 Multidrug Resistan
32	1568	60.5	4378	11	AAQ04522 Rat multidrug resi
33	1567	60.4	4425	21	AAZ52048 Rat multidrug resi
34	1566.5	60.4	4369	21	AAZ52047 Sequence of human
35	1564.5	60.4	4669	14	AAQ52726 Human polynucleoti
36	1564	60.3	3840	24	ABL91687 Human MDR-1 DNA.
37	1562.5	60.3	3988	21	AAZ88973 Murine multidrug r
38	1562.5	60.3	3988	21	AAZ88973 Murine BCRP DNA rel
39	1559.5	60.2	4788	24	ABA94365 Human ATP binding
40	1559.5	60.2	4788	21	AAZ94742 Human MDR-3 DNA.
41	1479	57.1	3924	21	AAZ94742 Gene #2299 used to
42	1479	57.1	3924	24	ABK63653 Rat sequence diffe
43	1479	57.1	3924	24	ABK63653 Human ATP binding
44	1445	55.7	3912	24	AAZ94744
45	1273.5	49.1	4776	21	AAZ94744

ALIGNMENTS

RESULT 1
ABK83223
ID ABK83223 standard; cDNA; 3699 BP.
XX
AC ABK83223;
XX
DT 27-AUG-2002 (first entry)
XX
DE Human transporter and ion channel, TRICH14, Incyte ID 7472030CB1, cDNA.
XX
KW Human; ss; gene; transporter and ion channel; TRICH; transport disorder;
KW neurological disorder; muscle disorder; immunological disorder; cancer;
KW scleroderma; systemic lupus erythematosus; allergy; leukaemia;
KW cell proliferative disorder; cervical cancer; breast cancer;
KW neurodegenerative disorder; Parkinson's disease; Alzheimer's disease;
KW myotonic dystrophy; catatonia; endocrine disorder; diabetes;
KW Grave's disease; gastrointestinal disorder; Crohn's disease;
KW renal disorder; Good pasture's syndrome; viral infection; cirrhosis;

KW bacterial infection; fungal infection; parasitic infection;
 KW protozoal infection; helminthic infection; cardiovascular disorder;
 KW atherosclerosis; hepatic disease.

XX Homo sapiens.

XX WO200240541-A2.

XX 23-MAY-2002.

XX 25-OCT-2001; 2001WO-US46055.

XX 27-OCT-2000; 2000US-243989P.

XX 03-NOV-2000; 2000US-245904P.

XX 09-NOV-2000; 2000US-247673P.

XX 17-NOV-2000; 2000US-249661P.

XX 20-NOV-2000; 2000US-252232P.

XX 01-DEC-2000; 2000US-250790P.

XX (INCYTE) INCYTE GENOMICS INC.

XX Tang YT, Yue H, Nguyen DB, Hafalia AJA, Elliott VS, Lu Y;

XX Wallia NK, Yao MG, Baughn WR, Gandhi AR, Ding L, Sanjanwala M;

XX Ramkumar J, Arvizu C, Gietzen KJ, Lal PG, Azimzai Y, Khan FA;

XX Thangavelu K, Thornton M, Lu DAM, Tribouley CM, Warren BA;

XX Ison CH, Das D, Raumann BE, Policky JL, Kearney L;

XX WPI; 2002-463570/49.

XX P-PSDB; ABG61544.

XX New transporters and ion channels (TRICH) polypeptides, useful for

XX diagnosing, preventing, and treating disorders associated with an

XX abnormal expression or activity of TRICH, e.g. immunological, muscular

XX or renal disorders

XX Claim 5; Page 172; 178pp; English.

XX The invention relates to human transporters and ion channels (TRICH)
 CC polypeptides, a naturally occurring amino acid sequence 90 % identical to
 CC TRICH, a biologically active fragment of TRICH or an immunogenic fragment
 CC of TRICH. Also included are an isolated polynucleotide encoding TRICH,
 CC a recombinant polynucleotide comprising a promoter sequence operably
 CC linked to the TRICH polynucleotide, a cell transformed with the
 CC recombinant polynucleotide, a transgenic organism comprising the
 CC recombinant polynucleotide, an isolated antibody that binds specifically
 CC to TRICH, and screening for compounds which bind to TRICH, modulate
 CC TRICH, modulate TRICH expression or are ant/agonists of TRICH.

CC The polypeptides are useful for diagnosing, treating, and
 CC preventing transport, neurological, muscle, immunological disorders
 CC (e.g. scleroderma, systemic lupus erythematosus, allergies), cell
 CC proliferative disorders such as cancers (e.g. leukaemia, cervical or
 CC breast cancers), neurodegenerative disorders (e.g. Parkinson's disease,
 CC Alzheimer's disease), muscular disorders (e.g. myotonic dystrophy,
 CC catatonla), endocrine disorders (e.g. diabetes, Grave's disease),
 CC gastrointestinal disorders (e.g. Crohn's disease), renal disorders
 CC (e.g. Good pasture's syndrome), viral, bacterial, fungal, parasitic,
 CC protozoal and helminthic infections, cardiovascular disorders (e.g.
 CC atherosclerosis), or hepatic diseases (e.g. cirrhosis) and many
 CC other diseases and disorders detailed in the specification. They can also
 CC be used in assessing the effects of exogenous compounds on the
 CC expression of nucleic acid and amino acid sequences of transporters and
 CC ion channels. TRICH or its fragments may also be used in screening for
 CC compounds that specifically bind to and modulate the activity of TRICH.
 CC The polynucleotides can be used to create knock-in humanised animals or
 CC transgenic animals to model human disease. The present sequence
 CC encodes a TRICH protein.

XX Sequence 3699 BP; 1116 A; 707 C; 860 G; 1016 T; 0 other;

XX Alignment Scores:

Pred. No.: 1.53e-246 Length: 3699

Score: 2460.50 Matches: 495

Percent Similarity: 95.40% Conservative: 3

Best Local Similarity: 94.83% Mismatches: 9
 Query Match: 94.93% Indels: 15
 DB: 24 Gaps: 2

US-09-873-409-8 (1-514) x ABK83223 (1-3699)

QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20

Db 145 ATGATCCTGGGTATACCTGACATCAGTCTTCAATGGAGCCTCCCTTTAATGCCACTG 204

QY 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40

Db 205 TGTATAGGAGAAATAGTATACCTTATTAGTGGATGTCTAGTCCACACTAACACAACA 264

QY 41 SerPhePheArg-----LeuThr 46

Db 265 AATTATCAGAACTGTACTCAGTCTCAAGAGAGCTGAATGAAGATATGATCTGTTGACC 324

QY 47 LeuTyrTyrValGlyIleGlyValAlaLeuIlePheGlyTyrIleGlnIleSerLeu 66

Db 325 CTGTATTATTGTTGGAATAGTGTGCTGCTTGGTTTGGTTTACATACAGATTTCCTTG 384

QY 67 TrpIleIleThrAlaAlaArgGlnThrLysArgIleAargLysGlnPhePheHisSerVal 86

Db 385 TGGATTATTAACCTGCAGCAGCAGACAGAGAGATTTCGAAACAGTTTTTTCATTGAGTT 444

QY 87 LeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGluLeuAsnThrArgMet 106

Db 445 TTGGCAGACAGACATCGCTGGTTTGGTTTGGTTTGGTTTGGTTTGGTTTGGTTTGGTTT 504

QY 107 Thr---AspIleAspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGln 125

Db 505 ACAGATGACATTGACAAATCAGTGTGTTGGAGATAAGATTGCTCTGTTGTTTCAA 564

QY 126 AsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLysGlyTrpLysLeuThr 145

Db 565 AACATGCTACTTTTTTCGATTGGCTGGCAGTTGGTTGGTTGGTTGGTTGGTTGGTTGGTT 624

QY 146 LeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAlaCysSerArgMet 165

Db 625 CTAGTCACTCTATCCAGTCTCTCTTATTAATGGCTTCAGCGCAGCATGTTCTAGATG 684

QY 166 ValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGlu 185

Db 685 GTCATCTCATTTGACAGTAAGGAATTAAGTGCCTATTCCAAAGCTGGGGCTGGCAGAA 744

QY 186 GluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGlnLysGluLeuGln 205

Db 745 GAAAGTCTTGTCATCCAGACAGTCATAGCCTTTAGGGCCCGAGAGAAAGAACTTCAA 804

QY 206 AtgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSer 225

Db 805 AGGTATACAGAGTCTCAAGATGCAAGAGATTTTGGCATATAAAGAGACTATAGCTTCA 864

QY 226 LysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrp 245

Db 865 AAAGTGTCCTTGGTCTGTGTACTTCTTTATGAATGAACCTTATGAGACTTGTCTTTTGG 924

QY 246 TyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAla 265

Db 925 TATGGAACTCTCTTGTATCTTATGGAGAACCTTGGATATACCATCGGGAGCTGTCTTGTCT 984

QY 266 ValPhePheSerValIleHisSerSerTyrCysIleGlyValAlaValProHisPheGlu 285

Db 985 GTTTCTTTTAGTGTATTCATAGTATTGATTTGGAGCAGCTCCCTCTACTTTGAA 1044

QY 286 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPro 305

Db 1045 ACCTTCGCAATACCCCGAGGAGCTGCTTTCATATTTTCCAGGTATTATGATAAGAACCC 1104

QY 306 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 325

Db 1105 AGTATAGTAACTTTTCCACAGCTGGATATAAACCTGAATCCATAGAGGAACCTGTGGAA 1164

Qy	326	PhcIyAaSnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu	345
Db	1165	TTTTAAAAATGTTTCTTTCAATTATCCATCAAGGCCCATCTATCAAGATTCTGAAGAGTCTG	1224
Qy	346	AsnLeuAArgIleLysSerGlyGluThrValAlaLeuValGlyLysLeuAenGlySerGlyLys	365
Db	1225	AACTCTCGGAATTAAAGTCTGGAGAGACAGTCGCTCTGGTCGGTCTCAATGGCAGTGGGAAG	1284
Qy	366	SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal	385
Db	1285	AGTACGGTAGTCCAGCTTCTGCAGAGATTATATGATCCGGATGATGGCTTTATCATGTG	1344
Qy	386	AspGluAenAspIleAArgAlaLeuAenValAArgHisTyrAArgAspHisIleGlyValVal	405
Db	1345	GATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATTCGAGACCATATTGGAGTGGTT	1404
Qy	406	SerGlnGluProValLeuPheGlyThrThrIleSerAenAsnIleLysTyrGlyArgAsp	425
Db	1405	AGTCAGAGAGCTGTTTGTTCGGAGCCACCATCAGTAACAATATCAAGTATGGACGAGAT	1464
Qy	426	AspValThrAspGluMetGluAArgAlaAlaAArgGluAlaAAsnAlaTyrAspPheIle	445
Db	1465	GATGTGACTGATCAAGAGATGGAGAGAGCAGCAAGGGAAGCAATGCGTATGATTTATC	1524
Qy	446	MetGluPheProAenLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGly	465
Db	1525	ATGAGTTCCTTAATAAATTAATCATTTGGTAGGGAAAAAAGAGAGCTCAAAATCAGTGGGA	1584
Qy	466	GlyGlnLysGlnAArgIleAlaIleAlaAArgAlaLeuValAArgAsnProLysIleLeuIle	485
Db	1585	GGCGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTAGTTCGAAAACCCCAAGATTTCTGATT	1644
Qy	486	LeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeu	505
Db	1645	TTAGATGAGCTACGTCTGCCCTGGATTTCAGAAACCAAGTCAGCTGTTCAGAGTGCACATG	1704
Qy	506	GluLys 507	
Db	1705	GAGAAG 1710	
RESULT 2			
AAZ49334			
ID	AAZ49334	standard; cdna; 4189 BP.	
XX			
AC	AAZ49334;		
XX			
DT	14-MAR-2000	(first entry)	
XX			
DE	Murine multidrug resistance-1 (MDR-1) cdna.		
XX			
KW	Multidrug resistance; MDR-1; P-glycoprotein;		
KW	transmembrane efflux pump; haematopoietic stem cell; transduction;		
KW	bone marrow transplantation; chemotherapy; radiation therapy; cancer;		
KW	gene therapy; gene replacement; genetic defect; thalassaemia;		
KW	Gaucher's disease; sickle cell anaemia; leukaemia; ex vivo expansion;		
KW	cytokine; ds.		
OS	Mus sp.		
XX			
Key	Location/Qualifiers		
FF	1..3831		
CDS	/*tag= a		
FT	/product= "Murine MDR-1 protein"		
FT			
XX			
XX	WO9961589-A2.		
XX			
XX	02-DEC-1999.		
XX			
XX	27-MAY-1999;	99WO-US11825.	
XX			
XX	28-MAY-1998;	98US-0086988.	
XX			

(STUD.-) ST JUDE CHILDREN'S RES HOSPITAL.

XX Sorrentino B, Bunting K;
XX
XX WPI; 2000-072615/06.
XX P-PSDB; AAY58188.
XX
XX Ex vivo expansion of hematopoietic stem cells transduced with a
XX sequence encoding human multidrug resistance-1, used for bone marrow
XX transplantation -
XX
XX Disclosure; Page 90-93; 113pp; English.
XX
XX This sequence represents cDNA encoding murine multidrug
XX resistance protein MDR-1. MDR-1 is a transmembrane
XX efflux pump, responsible for the export of drugs from cells,
XX particularly cancer cells. The invention relates to transducing
XX hematopoietic stem cells with nucleic acid encoding an MDR protein
XX and culturing the modified cells. The modified haematopoietic stem
XX cells are useful in bone marrow transplantation (to reconstitute
XX hematopoietic systems in patients who have undergone chemotherapy or
XX radiation therapy) and in ex vivo gene therapy of genetic defects in
XX cells derived from hematopoietic stem cells, e.g. thalassemia,
XX Gaucher's disease, sickle cell anaemia or leukaemia. The modified
XX cells can also be used to identify factors involved in regulating
XX proliferation and differentiation in haematopoietic stem cells.
XX Haematopoietic stem cells that express MDR-1 will be protected against
XX chemotherapeutic agents, so can be engrafted while the patient is
XX undergoing chemotherapy. Expansion of (rare) haematopoietic stem cell
XX provides sufficient cells to permit standard biochemical analysis.
XX Overexpression of MDR-1 allows cytokine-driven expansion of
XX haematopoietic stem cells by at least 10-fold compared with a maximum
XX of 4-fold in known procedures.
XX
XX Sequence 4189 BP; 1204 A; 875 C; 1028 G; 1082 T; 0 other;
XX

Alignment Scores:
Pred. No.: 1.74e-155 Length: 4189
Score: 1590.00 Matches: 307
Percent Similarity: 76.32% Conservative: 99
Best Local Similarity: 57.71% Mismatches: 94
Query Match: 61.34% Indels: 32
DB: 21 Gaps: 5

US-09-873-409-8 (1-514) x AAZ49334 (1-4189)

Qy 1 MetileuGlylleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 148 ATGATTCGGAACTCTCGCTGCTATTATTCATGGAACATTACTTCCCTCTTGATGCTG 207
Qy 21 ValLeuGlyGluMetSerAspAsnLeulle----- 30
Db 208 GTGTTTGGAAACATGACAGATAGTTTACAAAGCAGACCGCAGTCTGGAGGAAGAG 267
Qy 31 -----SerGlyCysLeuValGlnThrAsnThr----- 39
Db 268 ACTAATCAAGTGGACCAACAGTACTCTGATCATCAGCAACAGCAGTCTGGAGGAAGAG 327
Qy 40 -----TyrSerPhePheArgLeuThrLeuTyrTyrValGlylleGlyValAlaLa 56
Db 328 ATGCCCATATACGCTTAC-----TATTACACCGGGATTGGTGTCTGTGTGTG 372
Qy 57 LeuilePheGlyTyrileGlnIleSerLeuTrrpilelthAlaAlaArgGlnThrLys 76
Db 373 CTCAATGTCCTACATCCAGGTTTCACTTTGGTGCTGGCAGCTGGAGACAGATACAC 432
Qy 77 ArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrrPheAspSer 96
Db 433 AAGATTAGGCAGAGTTTTTCCATGCTATAATCAATCAGGAGATAGCTGGTTTGATGTG 492
Qy 97 CysAspIleGlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGly 115
Db 493 CATGATGTGGGAGCTCAACACCCGCTCACAGATGATGCTTCCAAAATTAATGACGGA 552


```
DB: 24 Gaps: 5
US-09-873-409-8 (1-514) x ABA94367 (1-4189)
QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 148 ATGATTTCTGGAACTCTCGCTGCTATTATTCATGGAACATTACTTCCCTCTCTGATGCTG 207
QY 21 ValLeuGluMetSerAspAsnLeuIle----- 30
Db 208 GTGTTTGGAAACATGACAGATAGTTTACAAAAGCAGAACCCAGTATTCTGCCAAGCATT 267
QY 31 -----SerGlyCysLeuValGlnThrAsnThr----- 39
Db 268 ACTAATCAAAAGTGGACCCCAACAGTACTCTGATCATCAGCAACAGCAGTCTGGAGGAAGAG 327
QY 40 -----TyrSerPhePheArgLeuThrLeuTyrValGlyIleGlyValAlaAla 56
Db 328 ATGCCATATACCCCTAC-----TATTACACCGGATTTGGTGTGCTGTG 372
QY 57 LeuIlePheGlyTyrIleGlnIleSerLeuTyrIleIleThrAlaAlaArgGlnThrLys 76
Db 373 CTCATAGTTGCCCTACATCCAGGTTTCACTTTGGTGCCTGCGCAGCTGGAACAGATACAC 432
QY 77 ArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSer 96
Db 433 AAGATTAGGCAGGAAGTTTTCATGCTGCTATTAATGAATCAGGAGATAGGCTGGTTGATGTG 492
QY 97 CysAspIleGlyLeuLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGly 115
Db 493 CATGATGTTGGGAGCTCAACACCGGCTCACAGATGATCTCCAAAATTAATGACGGA 552
QY 116 IleGlyAspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAla 135
Db 553 ATTGGTGACAAATTGGAGTGTTCATCCATAACCAACATTTTAGCCGGTTTATC 612
QY 136 ValGlyLeuValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIle 155
Db 613 ATAGGATTTATAAGTGGTTGGAAGCTAACCCCTTGTCAATTTTGGTGTGTCAGCCCTCTTATT 672
QY 156 MetAlaSerAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSer 175
Db 673 GGATGTCATCTGCTTGTGGCAAGAGTATTGACTTCATTACTATAAGGAACCTCCAG 732
QY 176 AlaTyrSerLysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIle 195
Db 733 GCTATTGCAAAAGCTGGAGCAGTGTCTGAGAAAGTCTTAGCAGCCATCAGAACTGTGATT 792
QY 196 AlaPheArgAlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLys 215
Db 793 GCCTTTGGAGGACACAGAAAGAACTTGAAGGTACAAATAAAATTTAGAAGAAGCTAAA 852
QY 216 AspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePhe 235
Db 853 AATGTTGGCATAAAGAAAGCTATCACAGCCAGCATTTTCGATAGGCATTTGCTACCTGTTG 912
QY 236 MetAsnGlyThrTyrGlyLeuAlaPheTyrGlyThrSerLeuIleLeuAsnGlyGlu 255
Db 913 GTCTATGCATCATATGACATGGCATTCGTGTATGGGACATCCTTGGTCTCTCAATGAA 972
QY 256 ProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyr 275
Db 973 -----TATTCTATTGGAGAAGTGTCTTACTGTCTCTCTCTATTATTGTTGGGACTTTT 1026
QY 276 CysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPhe 295
Db 1027 AGTATTGGACATTTGGCCCAACATAGAACGCTTTTCCAAACGACGAGGGGAGCCCTTT 1086
QY 296 HisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyr 315
Db 1087 GAAATCTTCAAGATAATTGATAACGACGACAGCATTCGACGCTTCTCAACAAAGGGCTAC 1146
QY 316 LysProGluSerIleGlyThrValGluPheLysAsnValSerPheAsnTyrProSer 335
```

```
Db 1147 AAACACAGACAGTATATATGGAAACCTTAGAGTTTAAATAATGTTTCACTTCAACCTACCATCG 1206
QY 336 ArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrVal 355
Db 1207 AGAAGCGAAGTTTCAGATCTTGAAGGCGCTCAATCTGAAGGTGAAGAGCGACAGCGGTG 1266
QY 356 AlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeu 375
Db 1267 GCCTTGGTTGGCAACAGTGGCTGTGGAAAAAGCACAACTGTCTCCAGCTGATGCAGAGGCTC 1326
QY 376 TyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnVal 395
Db 1327 TAGACCCCTCGAGGCGTGGTTCAGTATCGACGGCAAGACATCAGAACCATCAATGTG 1386
QY 396 ArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThr 415
Db 1387 AGTATCTCAGGGAGATCATTTGGTGGTGTGAGTCAGAACTGTGCTGTTGCCACACG 1446
QY 416 IleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluMetGluArgAla 435
Db 1447 ATGCCGAGAACATTCGCTATGCCGCGAGAAGATGTCCCATGTGATGAGATTGAGAAAGCT 1506
QY 436 AlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeu 455
Db 1507 GTCAAGGAAGCCCAATGCCCTATGACTTCATCATGAAACTGCCCACTTGTGACCCCTG 1566
QY 456 ValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArg 475
Db 1567 GTTGGTGAAGAGGGCGCAGCTGAGTGGGGNACAGAAACAGAGATCGCCATGCGCCG 1626
QY 476 AlaLeuValArgAsnProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspSer 495
Db 1627 GCCTGCTGCGCAATCCCAAGATCCTTTTGTGGACGAGGCCACCTCAGCCCTGGATACA 1686
QY 496 GluSerLysSerAlaValGlnAlaAlaLeuGluLys 507
Db 1687 GAAAGTGAAGCTGTGTGCGGCCCGCCTGGATAAG 1722
RESULT 4
AAQ38950
ID AAQ38950 standard; DNA; 4313 BP.
XX
AC AAQ38950;
DT
XX 28-JUL-1993 (first entry)
DE Mouse multidrug resistance sequence.
XX
KW mdr gene; Lambda DR11 clone; ss.
XX
OS Mus musculus.
XX
FH Key Location/Qualifiers
FT CDS 110..3940
FT /*tag= a
FT /phenotype= multidrug_resistance
XX
PN US5198344-A.
XX
PD 30-MAR-1993.
XX
XX
XX 15-JUL-1986; 86US-0885951.
XX
XX 15-JUL-1986; 86US-0885951.
XX 06-FEB-1991; 91US-0652311.
XX
XX (MASI ) MASSACHUSETTS INST TECHNOLOGY.
XX
XX Croop JM, Gros P, Houseman DE;
XX
XX WPI; 1993-126077/15.
DR
P-PSDB; AAR35199.
```


XX DNA sequence which confers multi-drug resistance on sensitive
PT mammalian cells - used to preserve bone marrow cells during
PT chemotherapy to prevent infection
XX
XX Claim 1; Fig 8; 22pp; English.
XX
XX A cDNA library was constructed from mRNA which had been isolated
CC from a drug-sensitive mouse cell-line. Two mouse cDNA molecules
CC complementary to the mRNA species encoded by 2 related but distinct
CC mdr genes were isolated and cloned. One of the cDNA clones (lambda
CC dril) is a full-length cDNA clone for one member of the mdr gene
CC family. When incorporated into prokaryotic expression vector pDREX4
CC (which allows high levels of transcription of the cDNA when
CC introduced into mammalian cells), the clone was shown to confer the
CC multidrug resistance phenotype upon transfection into drug-sensitive
CC mammalian cells.
XX
SQ Sequence 4313 BP; 1241 A; 904 C; 1058 G; 1110 T; 0 other;

Alignment Scores:

Pred. No.:	1.81e-155	Length:	4313
Score:	1590.00	Matches:	307
Percent Similarity:	76.32%	Conservative:	99
Best Local Similarity:	57.71%	Mismatches:	94
Query Match:	61.34%	Indels:	32
DB:	14	Gaps:	5

US-09-873-409-8 (1-514) x AAQ38950 (1-4313)

QY 1 MetileuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
DB 257 ATGATTCTGGGAATCTCGTCTGCTATTATCCATGGAACATTCTCCCTCTTGATCGTG 316

QY 21 ValLeuGlyGluMetSerAspAsnLeuLeu----- 30
DB 317 GTGTTTGGAAACATGACAGATAGTATTTTAAAGAGCAGAGCCAGTATCTGCGCAAGCATT 376

QY 31 -----SerGlyCysLeuValGlnThrAsnThr----- 39
DB 377 ACTAATCAAGTGGACCCACAGACTCTGATCATGACGACACAGCAGCTGGAGGAAGAG 436

QY 40 -----TyrSerPhePheArgLeuThrLeuTyrValGlyIleGlyValAlaAla 56
DB 437 ATGCCCATATACGCTAC-----TATTACACCGGATGTGTGCTGTG 481

QY 57 LeuIlePheGlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLys 76
DB 482 CTCATAGTTGCCTACATCCAGGTTTCACTTGTGTCCTGGCAGCTGGAAGACAGATACAC 541

QY 77 ArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSer 96
DB 542 AAGATTAGGCAGAAAGTTTTCCATGCTATATGAATCAGGAGATAGCTGTGTTGATGTG 601

QY 97 CysAspIleGlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGly 115
DB 602 CATGATGTTGGGAGCTCAACACCGGCTCACAGATGATGCTCCAAAATTAATGACGGA 661

QY 116 IleGlyAspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAla 135
DB 662 ATTGCTGACAAAATGGGATGTTTTTTCAGTCCATAACCAATTTTAGCCGGTTTTATC 721

QY 136 ValGlyLeuValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIle 155
DB 722 ATAGATTTTATAAGTGTGTGGAGCTAACTTGTCTATTTTGGCTGTGACGCCCTCTTATT 781

QY 156 MetAlaSerAlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSer 175
DB 782 GGATTGTCATCTGTTGTGGGCAAGGATTTGACTTCACTTACTATATAGGAATCTCAG 841

QY 176 AlaTyrSerLysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIle 195
DB 842 GCTTATGCAAAAGCTGGAGCAGTGTGCTGAAGAAGTCTTAGCGCCATCAGAACTGTGATT 901

QY 196 AlaPheArgAlaGlnGlnLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLys 215
DB 902 GCCTTTGGAGGACAAAGAGGAACTTGAAGGTACATAAATAATTTAGAGAAGCTAAA 961

QY 216 AspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePhe 235
DB 962 AATGTTGGCATAAAGAAAGCTATACAGCCAGCATTTTCATAGGCATTCCTACCTGTTG 1021

QY 236 MetAsnGlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsnGlyGlu 255
DB 1022 GTCATATGCATCATATGCATTCGCACTTCTGATGGACATCTCTGTCTCTCAATGAA 1081

QY 256 ProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerTyr 275
DB 1082 -----TATCTATTCGAGAGTGTCTACTGCTCTCTCTCTATTTTGTGGGGACTTTT 1135

QY 276 CysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPhe 295
DB 1136 AGTATTGGACACTTGGCCCAACATAGAGCCTTTTGCAAACGACGAGGGGAGCTTTT 1195

QY 296 HisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyr 315
DB 1196 GAAATCTTCAAGATAATTGATAACGAGCCAAAGCAATTTGACAGCTTCTCAACAAAGGGCTAC 1255

QY 316 LysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSer 335
DB 1256 AAACGACAGATATATATGGAACCTTAGAGTTTAAANAATGTTCACTTCACTACCATCG 1315

QY 336 ArgProSerIleLysIleLysGlyLeuAsnLeuArgIleLysSerGlyGluThrVal 355
DB 1316 AGAAGCGAAATTCAGATCTTGAAGGCGCTCAATCTGAAGGTGAAGAGCGGACAGCGTG 1375

QY 356 AlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuGlnArgLeu 375
DB 1376 GCCTTGTGTTGGCAACAGTGGCTGTGGAAAAGCACAACTGTCCAGCTGATGCGAGAGCTC 1435

QY 376 TyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnVal 395
DB 1436 TAGACCCCTCGAGGGCGGTGTCAGTATCGACGGACAGACATCAGACCATCAATGTG 1495

QY 396 ArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThr 415
DB 1496 AGTATCTGAGGAGATCATTTGTGTGTGTGAGTCAGGAACCTGTGTGTTGCCACCG 1555

QY 416 IleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAla 435
DB 1556 ATCGCCGAGAACATTCGCTATGCGGAGAGATGTCCCATGATGATGATTGAGAAGCT 1615

QY 436 AlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeu 455
DB 1616 GTCAGGAAGCCAAATGCTTATGATCTTCACTGAAACTGCCCCACCAATTTGACCCCTG 1675

QY 456 ValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArg 475
DB 1676 GTTGTGTGAGAGGGGCGCAGCTGAGTGGGGGACAGAAACAGAGAATCGCAITTCGCCGG 1735

QY 476 AlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSer 495
DB 1736 GCCTGTGTCGCAATCCCAAGATCCTTTTGTGGAGAGGCCACCTCAGCCCTGGATACA 1795

QY 496 GluSerLysSerAlaValGlnAlaAlaLeuGluLys 507
DB 1796 GAAAGTGAAGCTGTGTGTCAGGCGGCACCTGGATAG 1831

RESULT 5
AA03506
ID AA03506 standard; cDNA; 4279 BP.
XX
AC AA03506;
XX
DT 13-JUN-2001 (first entry)
XX

DE Dog P-glycoprotein (PGP) allelic variant (Genotype D) cDNA.
 KW Dog; P-glycoprotein allelic variant; PGP; multidrug transporter;
 KW MDRI; drug bioavailability; transgenic animal; genetic model, ss.
 XX
 OS Canis familiaris.
 XX
 XX Location/Qualifiers
 FT CDS 17..3862
 FT /tag= a
 FT /product= "Dog P-glycoprotein (PGP) allelic variant
 FT (Genotype D) protein"
 FT replace (91, T)
 FT /tag= b
 FT replace (607, C)
 FT /tag= c
 FT replace (1001, T)
 FT /tag= c
 FT replace (3458, A)
 FT /tag= c
 XX WO200123540-A2.
 XX
 XX 05-APR-2001.
 XX
 XX 28-SEP-2000; 2000WO-US26767.
 XX
 XX 28-SEP-1999; 99US-0156510.
 XX
 XX (GENT-) GENTEST CORP.
 PA
 XX
 PI Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;
 XX
 XX WPI: 2001-235373/24.
 DR P-PSDB; ABE00310.
 DR
 XX
 PT New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
 PT for determining the bioavailability of drugs and for screening for dog
 PT PGP inhibitors -
 XX
 PS Claim 9; Page 102-107; 11pp; English.
 XX
 CC The invention relates to dog P-glycoprotein (PGP) also referred
 CC as multidrug transporter (MDRI) and nucleic acids encoding them.
 CC The invention also includes fragments and biologically functional
 CC variants of dog P-glycoprotein, PGP and their nucleic acids are
 CC useful for determining the bioavailability of drugs and for
 CC screening PGP inhibitors. They are useful for the diagnosis and
 CC treatment of conditions characterised by PGP activity, by
 CC reducing or increasing PGP activity in a cell. PGP nucleic acids
 CC are used as oligonucleotide probes. Complements of PGP nucleic
 CC acids are useful as antisense oligonucleotides, to induce a PGP
 CC 'knockout' phenotype. They are used to prepare a non-human
 CC transgenic animal, which are valuable as genetic models for
 CC human diseases.
 CC The present sequence is dog P-glycoprotein (PGP) allelic variant
 CC (Genotype D) cDNA. The PGP enzyme functions as an efflux pump
 CC exporting small molecules across the cell membrane. This enzyme
 CC is a member of the ABC transporter family.
 XX
 SQ Sequence 4279 BP; 1296 A; 833 C; 1009 G; 1141 T; 0 other;
 Alignment Scores:
 Pred. No.: 5,96e-155 Length: 4279
 Score: 1585.00 Matches: 304
 Percent Similarity: 75.80% Conservative: 97
 Best Local Similarity: 57.47% Mismatches: 104
 Query Match: 61.15% Indels: 24
 DB: 22 Gaps: 4
 US-09-873-409-8 (1-514) x AAD03506 (1-4279)
 QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAlaCysLeuProLeuMetProLeu 20

Db 167 ATGTTGGTGGGACAAATGGCTGCATCATCCATGGAGCTGCACTCCCTCTCATGATGCTG 226
 QY 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu----- 34
 Db 227 GTTTTGGAAACATGCACAGATAGCTTTGCAAAATGCAGGAATTTCAAGAAACAAACTTTT 286
 QY 35 -----ValGlnThrAsnThrTyrSerPhePheArg----- 44
 Db 287 CCAGTTATAATTAATGAAAGTATTACGAACAATACACAACTTTTCATCAACCATCTGGAG 346
 QY 45 -----LeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePhe 59
 Db 347 GAGGAAATGACCACTATGCTTATTATTACAGTGGGATCGGTGGCGTGTGGTGGCT 406
 QY 60 GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg 79
 Db 407 GCTTACATCCAGTTTTCATCTGGTGCCTGGCAGCAGGAAGACAGATCTCAAAATTAGA 466
 QY 80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99
 Db 467 AAACAATTTTTTCATGCTATCATCGCAGGAGATTGGCTGTTTTCACGTGCATGACGTT 526
 QY 100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
 Db 527 GGGGAGCTTAACACCCGGCTCACAGCATGTCTCAAATCAATGAAGGAATTGGCGAC 586
 QY 119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
 Db 587 AAATTTGGNATGTTCTTTCAATCAATGACCAACATTTTTCACCGTTTATTAGTGGGTTT 646
 QY 139 ValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
 Db 647 ACACGTGGTGGAGCTAACCTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTTTCA 706
 QY 159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
 Db 707 GCGGCATCTGGGCAAGATATACTTCTCATTTACTGATAAAGAACTCTTGGCCTATGCA 766
 QY 179 LysAlaGlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArg 198
 Db 767 AAAGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGAACTGTGATTCCTTGA 826
 QY 199 AlaGlnGluLeuGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 218
 Db 827 GCACAAAAGAAAGAACTTGAAGGTACAAACAAAATTTAGAAAGCAAGCTAAAGAAATTGG 886
 QY 219 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPheMetAsnGly 238
 Db 887 ATAAAGAAAGCTATACGCCCAACATTTCTATTTGGTGGCGCTTTCTTATGTATCTATGCA 946
 QY 239 ThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 258
 Db 947 TCATATGCTCTGGCTTCTGGTATGGGACCTCTTGGTCTCTCCAGTGA-----TAT 1000
 QY 259 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 278
 Db 1001 ACTATTGGCAAGTACTCACTGCTTCTTTCTGTATTAAATTTGGGGCTTTTAGTATTGGA 1060
 QY 279 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 298
 Db 1061 CAGGCATCCCAAGCATTTGAAGCATTTGCAACGCAAGAGAGGAGCAGCTTATGAAATCTTC 1120
 QY 299 GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 318
 Db 1121 AAGATAATTGACAATAAACCAAGCATTTGACAGCTATTGAAAGAGTGACATAAACAGAT 1180
 QY 319 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 338
 Db 1181 AATATTAGGGAATTTGGAATTCAAATTTCACTTCAGTTACCTTCTCCAGTGA----- 1240
 QY 339 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 358


```
Db 707 GCGCCATCTGGCGCAAGATACCTATCTTCACTTAAAGAACTCTGGCCTATGCA 766
Qy 179 LyeAlaGlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArg 198
Db 767 AAAGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGAAGCTGTGATTGCTTTGGA 856
Qy 199 AlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 218
Db 827 GGCACAAAGAAGAAGCTTGAAGGTACACAAATAATTTAGAAGAAGCTAAGAAGATTGGG 886
Qy 219 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGly 238
Db 887 ATAAAGAAAGCTATACGCGCAACATTTCTATTGGTCCGCTTCTATTGATCATGCA 946
Qy 239 ThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 258
Db 947 TCATATGCTCTGGCTTCTCGTATGGGACCTCTTGGTCTCTCCAGTGAA-----TAT 1000
Qy 259 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 278
Db 1001 TCTATTGGACAAGTACTCACTGCTCTCTTTCTGTATTAAATGGGGCTTTTAGTATTGGA 1060
Qy 279 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaPheHisIlePhe 298
Db 1061 CAGCATCCCAAGCATTTGAAGCATTTGCAACGCAAGAGAGCAGCTTATGAATCTTC 1120
Qy 299 GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 318
Db 1121 AAGATAATTCACATAAACCAGCATTTGACAGCTATTGGAAGAGTGACATAAACAGAT 1180
Qy 319 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 338
Db 1181 AATATTAAAGGGAATTTGGAATTCAAAATGTTCACTTCACTTACCTTCTCGAAAGAA 1240
Qy 339 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrAlaLeuVal 358
Db 1241 GTTAAGATCTTAAGGGTCTCAACCTGAAGTTCAGAGTGGGAGAGTGGCGCTGGTT 1300
Qy 359 GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro 378
Db 1301 GGGACAGTGGCTGGGGAAGAGACACGCGTGCAGCTGATGCAGAGCTCTATGACCCC 1360
Qy 379 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 398
Db 1361 ACAGATGGCATGCTGTATTGATGGACAGGACATTAGGACCAATAATGTAAGCATCTT 1420
Qy 399 ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn 418
Db 1421 CGGGAATTAATCTGGTGGTGGTGCAGAGCCCTGTGTGTTGCCACCATGATGCTGAA 1480
Qy 419 AsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGlu 438
Db 1481 ACATTCGCTATGGCGCGGAATATGTCACCATGATGATGAGATTGAGAAGCTGTTAGGAA 1540
Qy 439 AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 458
Db 1541 GCCAATGCCTATGATTTTATCATGAACACTACCTAATAAATTTGACACTCTGGTTGGAGAG 1600
Qy 459 LysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal 478
Db 1601 AGAGGGGCCAGCTGAGTGTGTGACAGAAACAGAGAAATCGCCATTGCTGGGCCCTGGTT 1660
Qy 479 ArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 498
Db 1661 CGCAACCCCAAGATTTCTTCTGCTGGATGAGGCAACGTCAGCTCTGGACACTGAAGTGAA 1720
Qy 499 SerAlaValGlnAlaAlaLeuGluLys 507
Db 1721 GCAGTGGTTCAGGTGGCCCTGGATAAG 1747
RESULT 7
AAD03505
ID AAD03505 standard; cDNA; 4279 BP.
```

```
XX AAD03505;
XX 13-JUN-2001 (first entry)
XX Dog P-glycoprotein (PGP) allelic variant (Genotype B) cDNA.
XX Dog; P-glycoprotein allelic variant; PGP; multidrug transporter;
XX MDRI; drug bioavailability; transgenic animal; genetic model; ss.
XX Canis familiaris.
XX Key Location/Qualifiers
FH CDS 17..3862
FT /*tag= a
FT /product= "Dog P-glycoprotein (PGP) allelic variant
FT (Genotype B) protein"
FT replace (91, T)
FT /*tag= b
FT replace (607, C)
FT /*tag= c
XX WO200123540-A2.
XX 05-APR-2001.
XX 28-SEP-2000; 2000WO-US26767.
XX 28-SEP-1999; 99US-0156510.
XX (GENT-) GENTEST CORP.
XX Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;
XX WPI: 2001-235373/24.
XX P-PSDB; AAE00309.
XX New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
XX for determining the bioavailability of drugs and for screening for dog
XX PGP inhibitors -
XX Claim 9; Page 93-99; 11pp; English.
XX The invention relates to dog P-glycoprotein (PGP) also referred
XX as multidrug transporter (MDRI) and nucleic acids encoding them.
XX The invention also includes fragments and biologically functional
XX variants of dog P-glycoprotein. PGP and their nucleic acids are
XX useful for determining the bioavailability of drugs and for
XX screening PGP inhibitors. They are useful for the diagnosis and
XX treatment of conditions characterised by PGP activity, by
XX reducing or increasing PGP activity in a cell. PGP nucleic acids
XX are used as oligonucleotide probes. Complements of PGP nucleic
XX acids are useful as antisense oligonucleotides, to induce a PGP
XX 'knockout' phenotype. They are used to prepare a non-human
XX transgenic animal, which are valuable as genetic models for
XX human diseases.
XX The present sequence is dog P-glycoprotein (PGP) allelic variant
XX (Genotype B) cDNA. The PGP enzyme functions as an efflux pump
XX exporting small molecules across the cell membrane. This enzyme
XX is a member of the ABC transporter family.
XX Sequence 4279 BP; 1296 A; 833 C; 1008 G; 1142 T; 0 other;
Alignment Scores:
Pred. No.: 1,56e-154 Length: 4279
Score: 1581.00 Matches: 303
Percent Similarity: 75.80% Conservative: 98
Best Local Similarity: 57.28% Mismatches: 104
Query Match: 61.00% Indels: 24
DB: 22 Gaps: 4
US-09-873-409-8 (1-514) x AAD03505 (1-4279)
```



```
QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 167 ATGTTGGTGGGACAATGGCTGCGCATCATCCATGAGCTCCCTCTCATGATGCTG 226
QY 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu----- 34
Db 227 GTTTTGGAAACATGACAGATAGCTTTGCCAAATGCAGGAATTTCAAGAAACAAACTTTT 286
QY 35 -----ValGlnThrAsnThrTyrSerPheArg----- 44
Db 287 CCAGTTAATAATGAAGATTATGCAACAATACACAACATTTCAACACCATCTGGAG 346
QY 45 -----LeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePhe 59
Db 347 GAGGAATGACCAGTATGCTATTATACAGTGGATCGGTGCGTGGCTGGCT 406
QY 60 GlyTyrIleGlnIleSerLeuTyrIleThrAlaAlaArgGlnThrLysArgIleArg 79
Db 407 GCTTACATCCAGGTTTCATTCTGCTGCTGGCAGCAGGAAGACAGATACTCAAAATTAGA 466
QY 80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIle 99
Db 467 AAACAATTTTTCAATGATATCCGACAGGAGATTGCTGTTGACGTGCGATGAGCTT 526
QY 100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
Db 527 GGGGAGCTTAACACCGGCTCACAGACGATGCTCTCCAAATCAATGAAGAAATTGGCGAC 586
QY 119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
Db 587 AAAATTGGAATGTTCTTTCAATCAATAGCAACATTTTTCACCGGTTTTTAGTGGGGTTT 646
QY 139 ValLysGlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
Db 647 ACAGTGGTGGAAAGCTAACCTTGTGATTTGGCCATCAGCCCTGTCTTGCACTTCA 706
QY 159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
Db 707 GCCGCCATCTGGGCAAGATACTATCTTCAATTTACTATAAAGAACTCTTGGCCTATGCA 766
QY 179 LysAlaGlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArg 198
Db 767 AAAGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGAACTGTGATGCCCTTTGGA 826
QY 199 AlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 218
Db 827 GGACAAAGAAAGAACTTGAAGGTGACAAACAAATTTAGAAAGAACTGAAGAAATTGGG 886
QY 219 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPheMetAsnGly 238
Db 887 ATAAGAAAGCTATCACGGCCAACTATCTATTGTCGCGCTTTCTTATTGATCTATGCA 946
QY 239 ThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 258
Db 947 TCATATGCTCTGGCTTCTGTTATGGGACCTCCCTGCTCTCCAGTGAA-----TAT 1000
QY 259 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 278
Db 1001 TCTATTGGACAAAGTACTACGTCTCTTCTGCTATTATTAATGGGGCTTTTAGTATGGA 1060
QY 279 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 298
Db 1061 CAGGCATCCCAAGCATTTGAAGCATTTGCAACCGCAAGAGAGAGCAGCTTATGAAATCTTC 1120
QY 299 GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 318
Db 1121 AAGATAATTGCAATAAACCAGCATTTGACAGCTATTTCGAAGAGTGACATAAACCCAGAT 1180
QY 319 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 338
Db 1181 AATATTAAAGGAAATTTGGAATTTCAAAATGTTCACTTACGTTACCTCTTCGGAAGAA 1240
QY 339 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 358
```

```
Db 1241 GTTAAAGATCTTAAAGGTTCTCAACCTGAAGGTTCAGATGGGCGAGCAGTGGCGCTGGTT 1300
QY 359 GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro 378
Db 1301 GGGAAACAGTGGCTGCGGGAAGAGCACACGCTGCAGCTGATGCAGAGGCTCTATGACCCC 1360
QY 379 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 398
Db 1361 ACAGATGGCATGCTCTGATTGATGACAGGACATTTAGGACCATAATGTAAAGCATCTT 1420
QY 399 ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn 418
Db 1421 CGGAAATTAATCTGCTGCTGAGTCAGGAGCCTGTGTTGTTGCCACCATAGCTGAA 1480
QY 419 AsnIleLysTyrClyArgAspValThrAspGluGluMetGluArgAlaAlaArgGlu 438
Db 1481 AACATTCGCTATGCGCGCAAAATGTCACATGATGAGATGAGAAAGCTGTAAAGGAA 1540
QY 439 AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 458
Db 1541 GCCAATGCCATGATTTTATCATGAACCTACCTAATAAATTTGACACTCTGTTGGAGAG 1600
QY 459 LysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal 478
Db 1601 AGAGGGCGCCAGCTGAGTGTGTGGACAGAAACAGAGAATGCCCATTCGCGGCCCTGGTT 1660
QY 479 ArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 498
Db 1661 CGCAACCCCAAGATTTCTTCTGCTGATGAGGCAACGTGAGCTCTGGACACTGGAAGTGA 1720
QY 499 SerAlaValGlnAlaLeuGluLys 507
Db 1721 GCAGTGTTCAGTGGCGCTGGATAAG 1747

RESULT 8
AAF86128
ID AAF86128 standard; cDNA; 4195 BP.
XX AAF86128;
AC AAF86128;
XX
DT 25-JUN-2001 (first entry)
XX
DE Cynomologous monkey P-glycoprotein cDNA variant 1.
XX
KW Cynomologous monkey; P-glycoprotein; PGP; multidrug transporter; MDRI;
KW efflux pump; ss.
XX
OS Macaca fascicularis.
XX
FH Key Location/Qualifiers
FT CDS 100..3951
FT /tag= a
FT /product= "PGP"
FT /note= "P-glycoprotein"
FT misc_feature 376..384
FT /tag= b
FT /note= "Insertion of 9 nucleotides relative to PGP
FT allelic variant AAF86127"
XX
PN WO200123565-A1.
XX
PD 05-APR-2001.
XX
PF 28-SEP-2000; 2000WO-US26592.
XX
PR 28-SEP-1999; 99US-0156921.
PR 12-OCT-1999; 99US-0158818.
XX
PA (GENT-) GENTEST CORP.
XX
PI Stocker PJ, Steimel-Crespi DT, Crespi CL;
XX
```


DR WPI; 2001-316136/33.
 XX P-PSDB; AAB81065.
 PT Novel isolated nucleic acid encoding cynomologous monkey P-glycoprotein
 PT (PGP) and homologous PGP polypeptides are useful for predicting
 PT bioavailability of compound and increasing PGP transporter activity in
 PT cell
 XX
 PS Example 1; Page 59-65; 84pp; English.
 XX This invention relates to a polynucleotide sequence encoding a
 CC cynomologous monkey P-glycoprotein (PGP), and an allelic variant of the
 CC PGP protein. PGP, also known as multidrug transporter, MDR1 is a member
 CC of the ABC transporter superfamily. The enzyme serves as an efflux pump
 CC exporting small molecules across the cell membrane. The invention
 CC includes a cynomologous monkey (Macaca fascicularis) PGP coding sequence
 CC and protein, and also that of an allelic variant. The PGP polynucleotide
 CC sequence is useful for increasing PGP transporter activity in a cell.
 CC Antisense sequences of the cDNA are useful for inhibiting PGP transport
 CC activity in a mammalian cell. They may also be used for increasing the
 CC bioavailability of a drug. The present sequence represents cDNA encoding
 CC cynomologous monkey P-glycoprotein. This sequence contains a 9 nucleotide
 CC insert compared to the PGP allelic variant given in AAF86127.

XX Sequence 4195 BP; 1230 A; 801 C; 1039 G; 1125 T; 0 other;

Alignment Scores:

Pred. No.: 1-71e-154 Length: 4195
 Score: 1580.50 Matches: 306
 Percent Similarity: 75.85% Conservative: 96
 Best Local Similarity: 57.74% Mismatches: 103
 Query Match: 60.98% Indels: 25
 DB: 22 Gaps: 4

US-09-873-409-8 (1-514) x AAF86128 (1-4195)

QY 1 MetIleLeuGlyLeuLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
 DB 250 ATGTGGTGGGAACCTTTGGCTGGCCATCATCGAGCTGGACTTCCTCTCATGATGCTG 309
 QY 21 ValLeuGlyGluMetSerAspAsnLeuLeuSer-----GlyCys 33
 DB 310 GTGTTGGAGACATGACGGATACCTTTGCAATGTCAGGAAATTTAGGAGACT 369
 QY 34 LeuValGlnThrAsnThrTySer----- 41
 DB 370 CTGTTGTTTAAACACACATAATAGCAGTAATATCATCTATACGTGCGCGTCATGAATCTG 429
 QY 42 -----PhePheArgLeuThrLeuTyTyValGlyIleGlyValAlaAlaLeuLeu 58
 DB 430 GAGGAAGATATGACCGGTATGCCATTATATACAGTGGAAATTTGGTGGGTGCTGGTT 489
 QY 59 PheGlyTyIleGlnIleSerLeuTyIleThrAlaAlaArgGlnThrLysArgIle 78
 DB 490 GCTGCTTACATTCAGGTTTCATTTGGTGCTGCGACCTGGAAAGACAAATACACAAAT 549
 QY 79 ArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTyPheAspSerCysAsp 98
 DB 550 AGAAACAGTTTTTTCATGCTAATCCGACGAGGATAGCTGGTTGATGTGCACAT 609
 QY 99 IleGlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGly 117
 DB 610 GTTGGGAGCTTAACACCGGCTTACAGATGATGCTCTCCAAGATTAATGAAGAAATGGT 669
 QY 118 AspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGly 137
 DB 670 GACAAATTTGGAATGTTCTTTCAGTCAATGGCAACATTTTTTCACCTGGGTTTATAGTAGGA 729
 QY 138 LeuValLysGlyTyPheLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAla 157
 DB 730 TTTACAGCTGTGGAGCTGAACCTTGTGATTTGGCCATCATGCTCTCTTTGGACTG 789
 QY 158 SerAlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTy 177

DB 790 TCAGTGCAGTCTGGCAAGATACTGTCTTCTATTCTGATAAAGAACTCTTAGCTTAT 849
 QY 178 SerLysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPhe 197
 DB 850 GCANAGCTGGAGCAGTAGCTGAAGAGGCTTGGCAGCAATTAGAACTGTGATGCTAT 909
 QY 198 ArgAlaGlnGlyLysGluLeuGlnArgTyThrGlnAsnLeuLysAspAlaLysAspPhe 217
 DB 910 GGAGGACAAAGAAAGAACTCGAAAGGTACACAAATAATTTAGAAAGCTAAAGAAAT 969
 QY 218 GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyPhePheMetAsn 237
 DB 970 GGATAAAGAAAGCTATATACAGCCAAATATTTCTATAGGTGCTCTTCTCTGCTTATCTAT 1029
 QY 238 GlyTyThrGlyLeuAlaPheTyTyGlyTyThrSerLeuIleLeuAsnGlyGluProGly 257
 DB 1030 GCATCTTATGCTCTGGCCTTCTGGTATGGACCACTTGGTCTCTCAAGGAA----- 1083
 QY 258 TyThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyCysIle 277
 DB 1084 TATCTTATTTGGACAGTACTCACTGTATTTCTTCTGTATTAATTTGGGCTTTTAGTGT 1143
 QY 278 GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 297
 DB 1144 GGACGCGCATCTCCAAGCATTTGAAGCATTTGCAAAATGCAAGAGGACGCTTTTGAATC 1203
 QY 298 PheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyLysPro 317
 DB 1204 TTCAGATTAATGATTAATGAAGCCAGTATTTGACAGCTATTCGAGAGTGGGCAACCA 1263
 QY 318 GluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyProSerArgPro 337
 DB 1264 GATAATATTAAAGGAAATTTGGAATTCAGAAATGTTTCACTTCAGTTACCATCTCGAAA 1323
 QY 338 SerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu 357
 DB 1324 GAAGTTAAGATCTTGAAGGCGCTGAACCTGAAGGTGCAGAGTGGGAGCGTGGCCCTG 1383
 QY 358 ValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuGlnArgLeuTyAsp 377
 DB 1384 GTTGGAAACAGCGCTGCTGGGAAGACACACCGTCCAGCTGATGACAGAGCTTTATGAC 1443
 QY 378 ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis 397
 DB 1444 CCCACAGAGCGCATGCTCAGTGTGATGACAGGATATTAGGACCAATAACGTAAGGTT 1503
 QY 398 TyArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrIleSer 417
 DB 1504 CTACGGGAATCATCGGTGCTGAGTCAGGAACCTGTATTGTTTCCACACAGTAGCT 1563
 QY 418 AsnAsnIleLysTyTyGlyArgAspValThrAspGluGluMetGluArgAlaAlaArg 437
 DB 1564 GAAACATTCGCTATGCTGTAAGATGTCACCATGATGAGATTGAGAAAGCTGTGAC 1623
 QY 438 GluAlaAsnAlaTyTyAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly 457
 DB 1624 GAAGCCAAATGCTATGCTTATCATGAACTGCCTCAGAAATTTGCACACCTGGTTGA 1683
 QY 458 GluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaAlaLeu 477
 DB 1684 GAGAGAGGGGCCAGCTGAGTGGTGGGACAGAGCAGAGGATCGCCATTGCACGTGCCCTG 1743
 QY 478 ValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 497
 DB 1744 GTTCGCAACCCCAAGATCTCTGCTGGACGAGGCCACGTCAGCTTGGACACAGAAAGT 1803
 QY 498 LysSerAlaValGlnAlaAlaLeuGluLys 507
 DB 1804 GAAGCAGTGTTCAGGTGCTCTGGATAAG 1833
 RESULT 9
 AA249332

ID AAZ49332 standard; cDNA; 3860 BP.
 AC AAZ49332;
 XX 14-MAR-2000 (first entry)
 DT Human wild-type multidrug resistance-1 (MDR-1) cDNA.
 DE
 XX Multidrug resistance; MDR-1; P-glycoprotein;
 KW transmembrane efflux pump; haematopoietic stem cell; transduction;
 KW bone marrow transplantation; chemotherapy; radiation therapy; cancer;
 KW gene therapy; gene replacement; genetic defect; thalassaemia;
 KW Gaucher's disease; sickle cell anaemia; leukaemia; ex vivo expansion;
 KW cytokine; wild-type; ds.
 XX
 OS Homo sapiens.
 XX
 PH Key Location/Qualifiers
 FT CDS 1..3843
 FT /*tag= a
 FT /product= "Human wild-type MDR-1 protein"
 FT mutation replace (553..555, GTT)
 FT /*tag= b
 FT /note= "cDNA sequence of G185V human mutant MDR-1 given
 FT in AAZ49333"
 XX
 XX WO9961589-A2.
 XX
 XX 02-DEC-1999.
 XX
 XX 27-MAY-1999; 99WO-US11825.
 XX
 XX 28-MAY-1998; 98US-0086988.
 XX
 XX (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
 XX
 XX Sorrentino B, Bunting K;
 DR WPI; 2000-072615/06.
 DR P-PSDB; AAY58186.
 XX
 XX Ex vivo expansion of hematopoietic stem cells transduced with a
 PT sequence encoding human multidrug resistance-1, used for bone marrow
 PT transplantation -
 XX
 PS Claim 10; Page 68-70; 113pp; English.
 XX
 XX This sequence represents cDNA encoding human wild-type
 CC multidrug resistance protein MDR-1. MDR-1 is a transmembrane
 CC efflux pump, responsible for the export of drugs from cells,
 CC particularly cancer cells. Wild-type MDR-1 shows increased
 CC resistance to etoposide and decreased resistance to vinca
 CC alkaloids compared with a mutant form (AAV58187) where the Gly at
 CC position 185 is replaced by Val. The invention relates to transducing
 CC haematopoietic stem cells with nucleic acid encoding an MDR protein
 CC and culturing the modified cells. The modified haematopoietic stem
 CC cells are useful in bone marrow transplantation (to reconstitute
 CC haematopoietic systems in patients who have undergone chemotherapy or
 CC radiation therapy) and in ex vivo gene therapy of genetic defects in
 CC cells derived from haematopoietic stem cells, e.g., thalassaemia,
 CC Gaucher's disease, sickle cell anaemia or leukaemia. The modified
 CC cells can also be used to identify factors involved in regulating
 CC proliferation and differentiation in haematopoietic stem cells.
 CC Haematopoietic stem cells that express MDR-1 will be protected against
 CC chemotherapeutic agents, so can be engrafted while the patient is
 CC undergoing chemotherapy. Expansion of (rare) haematopoietic stem cells
 CC provides sufficient cells to permit standard biochemical analysis.
 CC Overexpression of MDR-1 allows cytokine-driven expansion of
 CC haematopoietic stem cells by at least 10-fold compared with a maximum
 CC of 4-fold in known procedures.
 XX
 XX Sequence 3860 BP; 1135 A; 746 C; 957 G; 1022 T; 0 other;

Alignment Scores:
 Pred. No.: 2,18e-154 Length: 3860
 Score: 1579.00 Matches: 306
 Percent Similarity: 77.04% Conservative: 100
 Best Local Similarity: 58.08% Mismatches: 99
 Query Match: 60.92% Indels: 22
 DB: 21 Gaps: 5
 US-09-873-409-8 (1-514) x AAZ49332 (1-3860)
 QY 1 MetIleLeuGlyIleLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
 DB 151 ATGGTGGGGAACTTGGCTGCCATCATCCATGGGGCTGGACTTCTCTCATGATGCTG 210
 QY 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
 DB 211 GTGTTGGGAATGACATATCTTGGCAATGCAGGAATTTAGAGATCTGATGTCA 270
 QY 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
 DB 271 AACATCACTAATAGAGTGTATATCAATGATACAGGCTTCTTCATGAATCTGGAGGAAGAC 330
 QY 44 -----ArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
 DB 331 ATGACCAGATATGCCCTATTATTACAGTGGAAATGGTGCTGGGGTGTCTGCTGTAC 390
 QY 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
 DB 391 ATTCAAGTTTCATTTGGTGCTGGCAGCTGGAGACAATAACAAATTTAGAAACAG 450
 QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
 DB 451 TTTTTCATGCTATATGCGACAGAGATAGGCTGCTTGTATGTGCACGATGTGGGAG 510
 QY 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
 DB 511 CTTAACACCCGACTTACAGATGATGTCTCTAGATTAATGAAGTATTGGTGCAAAAT 570
 QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
 DB 571 GGAATGTCTTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGATTACACGI 630
 QY 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaA 160
 DB 631 GGTTGGAAGCTAACCTTGTGATTTTGGCCATCAGTCTCTTCTTGGACTGTGAGTGTCT 690
 QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
 DB 691 GTCTGGGCAAGATACCTATCTTCTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 750
 QY 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
 DB 751 GGAGCAGTAGCTGAAGAGGCTTGGCAGCAATTAGAACCTGTGATTGCAITTTGAGAGCAA 810
 QY 201 GLuLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
 DB 811 AAGAAGAACTTGAAGGTACACAAAATTTAGAAAGAGCTAAAGAAATGGGATAAAG 870
 QY 221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
 DB 871 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTCTCTGCTGATCTATGATCTTAT 930
 QY 241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
 DB 931 GCTCTGCCCTTCGGTATGGGACCACTTGGTCTCTCTCAGGGGAA-----TATTCTATT 984
 QY 261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaA 280
 DB 985 GGACAAGTACTACTGTATCTTTCTGTATTAAATGGGGCTTTTAGTGTGGCAGGCA 1044
 QY 281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
 DB 1045 TCTCCAAGCATTTGAAGCAATTTGCAATGCAAGAGAGCAGCTTATGAAATCTTCAAGATA 1104


```
QY 301 IleAspGlySerProSerIleAspAsnPheserThrAlaGlyTyrLysProGluSerIle 320
Db 1105 ATTGATAATAAGCAAGTATTGACAGTATTGACAGTGGGCACAAACACAGATAATATT 1164
QY 321 GluGlyThrValGluPheLysAanValSerPheAsnTyrProSerArgProSerIleLys 340
Db 1165 AGGGAAATTTGGAAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAGAAGTTAAG 1224
QY 341 IleLeuLysGlyLeuAsnLeuAArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
Db 1225 ATCTTGAGGGCTTGAACTGAAGTGCAGAGTGGCAGACGGTGGCCCTGGTTGGAAC 1284
QY 361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
Db 1285 AGTGGCTGTGGGAAGAGCACACAGTCCAGCTGATGCAGAGGCTCTATGACCCACAGAG 1344
QY 381 GlyPheIleMetValAspGluAanAspIleAArgAlaLeuAanValArgHisTyrArgAsp 400
Db 1345 GGGATGGTCAGTGTGTATGGACAGGATATTAGGACCAATAATGTAAGGTTTCTACGGAA 1404
QY 401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
Db 1405 ATCATTTGGTGTGTGAGTACAGAACCTGTATTGTTTCCACACAGATAGCTGAACATT 1464
QY 421 LysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
Db 1465 CGCTATGGCCGTGAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCAAT 1524
QY 441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrIleValGlyGluLysGly 460
Db 1525 GCCTATGACCTTATCATGAACTGCCTCATAAATTTGACACCCCTGGTGGAGAGAGGG 1584
QY 461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db 1585 GCCCAGTTGAGTGGTGGCAGAGACAGAGATCGCCATTGCAGTCCCTGGTTGCGAAC 1644
QY 481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db 1645 CCCAGATCTCTCTGCTGGATGAGGCCACCTCAGCCTTGGACACAGAAAGCGAGCATG 1704
QY 501 ValGlnAlaAlaLeuGluLys 507
Db 1705 GTTCAGGTGGCTCTGGATAAG 1725

RESULT 10
ABA94365
ID ABA94365 standard; DNA; 3860 BP.
XX
AC ABA94365;
XX
XX
DT 26-MAR-2002 (first entry)
DE Human BCRP DNA related seq id No. 1.
XX
XX Stem cell; ATP transport protein; ATP-binding cassette; antiparkinsonian;
KW hepatotrophic; neurodegenerative; cytotstatic; antianemic; muscular; BCRP;
KW cardiant; gene therapy; ds.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FT 1..3843
CDS /*tag= a
FT
XX
XX WO200192877-A2.
PN
XX
XX 06-DEC-2001.
XX
XX 30-MAY-2001; 2001WO-US17459.
PF
XX
XX 31-MAY-2000; 2000US-0584586.
PR
XX
XX 29-MAY-2001; 2001US-0866866.
PR
```

```
XX
PA (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX Sorrentino B, Schuetz J;
XX
XX WPI: 2002-114368/15.
DR P-PSDB; ABB07266.
DR
XX
PT Identifying a stem cell, for treating e.g., muscular dystrophy,
PT myocardial infarction, Parkinson's disease, or neurodegenerative
PT disorders, comprises detecting the expression of an ATP transport
PT protein (BCRP) by a cell
XX
PS Disclosure; Page 53-55; 87pp; English.
XX
XX The invention provides a method of identifying and/or isolating a stem
CC cell that involves detecting the expression of an ATP transport protein
CC containing a conserved ATP-binding cassette (BCRP) by a cell in a sample
CC comprising stem cells. The isolated stem cells may be used in the
CC treatment of diseases such as muscular dystrophy, degenerative liver
CC disorder, myocardial infarction, Parkinson's disease, degenerative
CC disorders of the brain, and for tissue regeneration or replacement.
CC Haematopoietic cells can be used in bone marrow transplants (e.g., for
CC treatment of leukemia) and for ex vivo gene therapy for treating blood
CC diseases such as sickle cell anemia and thalassemia. The stem cells can
CC also be used as cell targets in gene therapy protocols. The present
CC sequence represents a sequence related to the BCRP for which no relevant
CC information has been provided in the specification.
XX
SQ Sequence 3860 BP; 1135 A; 746 C; 957 G; 1022 T; 0 other;

Alignment Scores:
Pred. No.: 2,18e-154 Length: 3860
Score: 1579.00 Matches: 306
Percent Similarity: 77.04% Conservative: 100
Best Local Similarity: 58.08% Mismatches: 99
Query Match: 60.92% Indels: 22
DB: 24 Gaps: 5

US-09-873-409-8 (1-514) x ABA94365 (1-3860)
QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 151 ATGGTGGTGGGAACCTTTGGCTGCCATCATCATCGGGCTGGAGACTCTCTCATGATGCTG 210
QY 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
Db 211 GTGTTTGGAGAAATGACAGATATCTTGGCAATGCAGGAATTTAGAGATCTGATGTCA 270
QY 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
Db 271 AACATCACTAATAGAAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 330
QY 44 -----ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db 331 ATGACCAAGATATCCCTATTATTACAGTGGAAATTTGGTGGGGTGGTGGTCTGCTGTAC 390
QY 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db 391 ATTCAGGTTTCATTTTGGTGGCTGGCAGCTGGGAAGACAAATACACAAATAGAAACAG 450
QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db 451 TTTTTCATGCTATATATCGCAGAGAGATAGGCTGGTTGATGTGCACGATGTTGGGAG 510
QY 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db 511 CTTAAACACCCGACTTACAGATGATGTCTCTAAGATTAAATCAGGTTATTGTCGACAAAT 570
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db 571 GGAATGTTCTTTCAGTCAATGGCAACATTTTCTACTGGGTTTATAGTAGGATTTACACGT 630
```


of mdrlb2 activity. It may also be used to establish assays to predict oral absorption and pharmacokinetics of drugs in humans, and thus enhance the design of formulations through the identification of compounds with optimal development characteristics (i.e., high oral bioavailability, uid (once a day) dosing, reduced drug interactions, reduced variability, and reduced food effects), specifically to avoid interactions with human mdrlb2. Transgenic and knockout animals created using DNA encoding the rat mdrlb2 may be used to gain an insight into treating and preventing human diseases such as cancer, inflammation, cardiovascular disease, central nervous system disorders, autoimmune disorders and kidney disease. The present sequence represents cDNA encoding rat mdrlb2.

SQ Sequence 4233 BP; 1210 A; 920 C; 1048 G; 1055 T; 0 other;

Alignment Scores:

Pred. No.: 5.12e-154 Length: 4233
Score: 1576.00 Matches: 308
Percent Similarity: 75.90% Conservative: 92
Best Local Similarity: 58.44% Mismatches: 105
Query Match: 60.80% Indels: 22
DB: 22 Gaps: 4

US-09-873-409-8 (1-514) x AAF27498 (1-4233)

Qy 1 MetileLeuGlyLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 170 ATGGCTCTGGGAACCTCTCGTGTATCATCCAGCAACCTCTGCTCCCTCTGATGCTG 229
Qy 21 ValLeuGlyGluMetSerAspAsnLeuLeu----- 30
Db 230 GTGTTGGATACATGACAGATAGTTTACCAAGCAGAGACCCGCAATTCGCCGAGGTT 289
Qy 31 -----SerGlyCysLeuValGlnThrAsnThrTyrSerPheArgLeu----- 45
Db 290 ACTAATCAAAGTGAATCAACAGTACACAGCCGTCAGCAGCAGCAGCTCTGGAGGAGAC 349
Qy 46 -----ThrLeuTyrTyrValGlyLeuGlyValAlaLeuLeuPheGlyTyr 61
Db 350 ATGGCCATGACGCTACTATTACCGGCAATGTCGCGGTGTCATCGTTCGCTAC 409
Qy 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrTyrArgIleArgLysGln 81
Db 410 ATCCAGTTTCACCTTGGTGGCTGGCAGCTGGGAGACAAATACACAAGATTAGGCAGAAG 469
Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db 470 TTTTTCATGCCATCATCAATCAGAGATAGGCTGGTTCAGCGTGAATGACGCTGGGGAG 529
Qy 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db 530 CTCACACCCGGCTCACAGATGAGCTCTCCAAATTAATGACGGAATTTGTGACAACTT 589
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db 590 GGAATGTCCTTTCAGTCCATAACGACATTTTCAGCCGGTTTATAATAGGATTTATAAGT 649
Qy 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db 650 GGTTCGAAGCTTAACCTTGTAATTTTGGCCGCTCAGCCCTCTTATTGGGTGTGTCATCTGCC 709
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysLeuLeuSerAlaTyrSerLysAla 180
Db 710 ATGTGGGCAAGGTAAGTACTTCTTACTTAATGAAGAACTCCAGGCTTATGCGAAAGCT 769
Qy 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db 770 GGACGAGTTGCCGAAGAAGTCTTAGCACCACATCAGACTGTGATGCTTTGGAGGACAA 829
Qy 201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
Db 830 AAGAAGGAACCTTGAAGAGTCAATAAATAATTTAGAAAGAGCTAAAGAGTTGGCATAAAG 889

Qy 221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
Db 890 AAAGCCATCAGCGCAACCATTTCCATAGGTATTCCTATGCTGTGCTATGCGTCTTAT 949
Qy 241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
Db 950 GCATGGCACTTCGTGTATGGGACCTCTTGGTCTCTCAATGAA-----TATCTATT 1003
Qy 261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
Db 1004 GGCAAGTCTTACCGTCTCTCTCTATTTATTTGGGACCTTTCAGTATTGGACATTTA 1063
Qy 281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
Db 1064 GCCCAACACATAGAGCCCTTTGCAATGCAAGAGGGGAGCCTTGAATCTTCAAGATA 1123
Qy 301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
Db 1124 ATTGATAATGAGCCCAAGCATTGACAGCTTCTCAACCAAGGGACACAAACGACAGATA 1183
Qy 321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
Db 1184 ATGGGAAATTTGGAATTTAAATGTTTACTTCACTACCATCACAAAGTGAAGTTAAG 1243
Qy 341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
Db 1244 ATCTTGAAGGGCTCAACCTGAAGGTGAAGAGCGGACAGCGTAGCCCTGGTTGGCAAC 1303
Qy 361 AsnGlySerGlyLysSerThrValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
Db 1304 AGTGGCTGTGGGAAAGCAACCTGTCCAGCTGCTGCAGAGGCTCTACGACCCCATAGAG 1363
Qy 381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
Db 1364 GCGCAGCTCAGTATCCAGGACAGGACATCAGACCATCAATGTGAGGTATCTGCCGGA 1423
Qy 401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
Db 1424 ATCATTTGGGTGTGAGTCAGGAACCCGCTGTTTGGCCACCACGATTCGCGAACAAT 1483
Qy 421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
Db 1484 CGTATGGCGGAGAAACCGTCAACCATGTGATGATAGAGAAAGCTCTCAAGGAAGCCAAT 1543
Qy 441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
Db 1544 GCCTATGACTTCAATGATAAATGCCCCCAAAATTTAACCCCTGGTGTGTGAGAGGG 1603
Qy 461 AlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db 1604 GCGCAGCTGAGTGGGGACAGAAACAGAGATCGCCATTTGCCCGGCGCTGTGTCGCAAC 1663
Qy 481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db 1664 CCCAAGATCCTTTTGTGTGATGAGGCGACCTGACGCTTGGACACAGAAAGCGCGGTG 1723
Qy 501 ValGlnAlaAlaLeuGluLys 507
Db 1724 GTTCAGCGCCCTCTGGATAG 1744
RESULT 13
AAD03488
ID AAD03488 standard; cDNA; 4279 BP.
XX
AC AAD03488;
XX
DT 13-JUN-2001 (first entry)
XX
DE Dog P-glycoprotein (PGP) cDNA #1.
XX
KW Dog; P-glycoprotein; PGP; multidrug transporter; MDR1;
XX drug bioavailability; transgenic animal; genetic model; ss.
XX

FT		/tag= b	
FT		/product= "wild-type Pgp"	
FT	3'UTR	4368..4669	
FT		/*tag= c	
XX			
XX	WO9821325-A1.		
PN			
XX			
PD	22-MAY-1998.		
XX			
XX	17-NOV-1997;	97WO-US21214.	
PF			
XX			
XX	15-NOV-1996;	96US-0752447.	
PR			
XX	(INGE-) INGENEX INC.		
PA	(UNII) UNIV ILLINOIS FOUND.		
FA			
XX			
PI	Mechetner E, Roninson IB;		
XX			
XX	WPI; 1998-297930/26.		
DR	P-PSDB; AAW48937.		
DR			
XX			
FT	Immunological reagent specific for P-glyco:protein - useful for		
PT	detecting multi-drug resistant cancer, isolating haematopoietic		
PT	cells and selective cell killing		
XX			
XX	Claim 1; Fig 1A; 89pp; English.		
PS			
XX			
CC	The present sequence represents the human P glycoprotein (Pgp)		
CC	cDNA which encodes the wild-type Pgp protein. Pgp is a transmembrane		
CC	efflux pump protein involved in multi-drug resistance of cancer cells		
CC	The invention provides methods for developing and using immunological		
CC	reagents specific for certain mutant forms of Pgp and wild-type Pgp i		
CC	a conformation associated with substrate binding or in the presence		
CC	of ATP depleting agents. An example of the immunological reagent		
CC	is the UIC2 monoclonal antibody (mAb).. mAB UIC2 specifically binds		
CC	to Pgp in a particular biochemical conformation and is capable of		
CC	inhibiting drug efflux from Pgp-expressing cells. The immunological		
CC	reagents are claimed to be useful for detecting Pgp expression in		
CC	mammalian cells, including low level expression, particularly in can		
CC	cells to diagnose multi-drug resistance. The invention claims that		
CC	these immunological reagents are more specific than known reagents fo		
CC	detecting Pgp and they also eliminate the need for costly and laborio		
CC	screening of Pgp inhibitors by cytotoxicity or dye exclusion methods.		
XX			
SQ	Sequence 4669 BP; 1393 A; 894 C; 1130 G; 1252 T; 0 other:		

[illegible]

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 30, 2003, 03:22:44 ; Search time 3535.38 Seconds
(without alignments)
2354.618 Million cell updates/sec

Title: US-09-873-409-8
Perfect score: 2592
Sequence: 1 MILGILASLVNGACILPLMLP.....SESKSAVQALEKDTPTVSP 514

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DEV=xlh

-O=/cgn2_1/USPTO spoil/US9873409/runat_27032003_115420_19240/app_query.fasta_1.7544
-DB=EST -QFMT=fastp -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-LOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US9873409 @CGN 1.1.30544 @runat_27032003_115420_19240 -NCPU=6 -ICPU=3
-NO XLPHY -NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DBLEXT=7

Database :

- 1: em_estba.*
- 2: em_esthum.*
- 3: em_estin.*
- 4: em_estmu.*
- 5: em_estov.*
- 6: em_estpl.*
- 7: em_estro.*
- 8: em_htc.*
- 9: gb_est1.*
- 10: gb_est2.*
- 11: gb_htc.*
- 12: gb_est3.*
- 13: gb_est4.*
- 14: gb_est5.*
- 15: em_estfun.*
- 16: em_estom.*
- 17: gb_gas.*
- 18: em_gas_hum.*
- 19: em_gas_inv.*
- 20: em_gas_pln.*
- 21: em_gas_vrt.*
- 22: em_gas_fun.*
- 23: em_gas_nam.*
- 24: em_gas_mus.*
- 25: em_gas_other.*
- 26: em_gas_pro.*
- 27: em_gas_rod.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1092	42.1	669	13	BM013981	BM013981 603639619
2	1068.5	41.2	1033	14	BM926413	BM926413 AGENCOURT
3	1029	39.7	760	13	BM016204	BM016204 603642659
4	926.5	35.7	929	14	BQ882401	BQ882401 AGENCOURT
5	918	35.4	545	9	AL040762	AL040762 DKF2P434C
6	856.5	33.0	948	13	BM471690	BM471690 AGENCOURT
7	750	28.9	894	12	BF584668	BF584668 602098406
8	744.5	28.7	872	11	AK020318	AK020318 Mus muscu
9	710	27.4	2676	11	AK014319	AK014319 Mus muscu
10	671.5	25.9	803	12	BG298756	BG298756 602396681
11	659	25.4	943	9	AL520322	AL520322 AL520322
12	654	25.2	405	9	AA243820	AA243820 zr67906.r
13	650	25.1	550	13	BJ075336	BJ075336 BJ075336
14	637	24.6	780	12	BG587938	BG587938 EST489713
15	625	24.1	559	12	BF692596	BF692596 602248949
16	621	24.0	606	14	BQ262340	BQ262340 fz85a08.x
17	621	24.0	681	14	BU006064	BU006064 QGG9N16.Y
18	604.5	23.3	932	17	AZ670821	AZ670821 ENTJN69TF
19	602.5	23.2	913	17	BH155700	BH155700 ENTRO54TR
20	599	23.1	880	17	AZ687805	AZ687805 ENTJUS2TF
21	596.5	23.0	871	17	AZ682350	AZ682350 ENTKB18TF
22	596.5	23.0	2564	11	AK020022	AK020022 Mus muscu
23	590.5	22.8	891	17	AZ682250	AZ682250 ENTKT68TR
24	585.5	22.6	853	17	AZ679807	AZ679807 ENTHI67TR
25	584.5	22.6	1813	11	U66688	U66688 Homo sapien
26	582.5	22.5	787	12	BG584063	BG584063 EST485823
27	581	22.4	586	13	BJ336461	BJ336461 BJ336461
28	579	22.3	661	12	BG303675	BG303675 fl36b04.x
29	578	22.3	575	10	AW342644	AW342644 fl81f06.x
30	577.5	22.3	886	17	AZ540627	AZ540627 ENTQ18TF
31	577	22.3	835	13	BJ349604	BJ349604 BJ349604
32	576.5	22.2	947	17	AZ683753	AZ683753 ENTIL96TF
33	576	22.2	547	13	BJ333166	BJ333166 BJ333166
34	574.5	22.2	885	17	AZ686798	AZ686798 ENTF28TR
35	574	22.1	542	13	BJ339421	BJ339421 BJ339421
36	572.5	22.1	860	17	AZ543293	AZ543293 ENTQ80TF
37	572.5	22.1	886	17	BH139685	BH139685 ENTNA47TR
38	571	22.0	715	14	BQ869512	BQ869512 QGD6G09.Y
39	570	22.0	814	13	BJ352304	BJ352304 BJ352304
40	570	22.0	820	13	BJ356540	BJ356540 BJ356540
41	570	22.0	1019	12	BG248052	BG248052 602359987
42	569	22.0	1341	11	AY108285	AY108285 Zea mays
43	565	21.8	657	14	BQ802058	BQ802058 WHE2821.F
44	564	21.8	636	13	BJ386538	BJ386538 BJ386538
45	564	21.8	750	13	BU445883	BU445883 BU445883

ALIGNMENTS

RESULT 1
BM013981
LOCUS 603639619F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:5415556 5',
DEFINITION mRNA sequence.
ACCESSION BM013981
VERSION BM013981.1 GI:16528335
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 669)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL
COMMENT

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12060 row: f column: 05
High quality sequence stop: 664.
Location/Qualifiers
1. .669
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5415556"
/clone_lib="NIH MGC 87"
/tissue_type="mammary adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: breast; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.383 Kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH MGC Library."

FEATURES

source

BASE COUNT 203 a 122 c 162 g 182 t
ORIGIN
Alignment Scores:
Pred. No.: 1,77e-119 Length: 669
Score: 1092.00 Matches: 215
Percent Similarity: 99.54% Conservative: 2
Best Local Similarity: 98.62% Mismatches: 0
Query Match: 42.13% Indels: 1
DB: 13 Gaps: 0
US-09-873-409-8 (1-514) x BM013981 (1-669)

QY 265 AlavalPheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPhe 284
Db 17 AGTGTTCCTTAATGTAATCAATAGCAGTATTGCATTGGAGCAGAGTCCCTCACTTT 76
QY 285 GluThrPheAlaIleAlaArgGlyValAlaPheHisIlePheGlnValIleAspLysLys 304
Db 77 GAACCTTCGCAATAGCCGAGGAGTCCCTTCATATTTTCCAGGTATTGATAAGAA 136
QY 305 ProSerIleAspAsnPheSerThrAlaGlyTyrIlysProGluSerIleGluGlyThrVal 324
Db 137 CCCAGTATAGATAACTTTTCCACAGCTGGATATAAACCTGAATCCATAGAAGAACTGTG 196
QY 325 GluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGly 344
Db 197 GAATTTAAAATGTTCTTTCAATATATCCATCAAGACCATCTCAAGATTCGAAGGT 256
QY 345 LeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGly 364
Db 257 CTGAATCTCAGAAATAGTCTGGAGACACAGTCGCTTGGTCTCAATGGAGTGGG 316
QY 365 LysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMet 384
Db 317 AAGAGTACGGTAGTCCAGCTCTGACAGGTTATATGATCGGATGATGCTTTATCATG 376
QY 385 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 404
Db 377 GTGGATGAGAAATGCAATCAGAGCTTTAAATGTGGCGCAATATCCAGACCATATTTGGAGTG 436
QY 405 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 424
Db 437 GTTAGTCAAGAGCCTGTTTTGTCGGACACCAATCATGTAACATATCAAGTATGACGA 496
QY 425 AspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 444
Db 497 GATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAAATCGGTATGATTTT 556

QY 445 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGlyAlaGlnMetSer 464
Db 557 ATCATGGAGTTTCCTAATAATTAATATACATTGGTAGGGAAAAAGAGCTCAATGAGT 616
QY 465 GlyGlyGlnLysGlnArg-IleAlaIleAlaArgAlaLeuValArgAsnPro 481
Db 617 GGAGGGCAGAAACAGAGGATCGCAATTGCTGCTTAGTTTGGAAACCCC 668

RESULT 2

BM926413
LOCUS
DEFINITION BM926413.1 GI:19376792
5', mRNA sequence.
ACCESSION BM926413
VERSION BM926413.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1033)
REFERENCE NIH-MGC http://mgi.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12818 row: o column: 22
High quality sequence stop: 3
High quality sequence stop: 686.
Location/Qualifiers
1..1033
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5764845"
/clone_lib="NIH MGC_114"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: EcoRV (destroyed); RNA source anonymous pool of 6
male brains, age range 23-27 yo. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.5 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 019. Note:
this is a NIH MGC Library."

FEATURES
source

BASE COUNT 289 a 212 c 218 g 312 t
ORIGIN
Alignment Scores:
Pred. No.: 2.3e-116 Length: 1033
Score: 1068.50 Matches: 229
Percent Similarity: 93.17% Conservative: 3
Best Local Similarity: 91.97% Mismatches: 17
Query Match: 41.22% Indels: 4
DB: 14 Gaps: 0
US-09-873-409-8 (1-514) x BM926413 (1-1033)
QY 266 ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGlu 285
Db 180 GTTTTCTTAGTGTAATCCATAGCAGTATTGCATTGGAGCAGAGTCCCTCACTTTGAA 239
QY 286 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysPro 305
Db 240 ACCTTCGAATAGCCCGAGGAGTGCCTTTTCATATTTTCCAGGTCAATGATAAGAAACCC 299


```
QY 306 SerIleAspAsnPheSerThrAlaGlyTyrIleProGluSerIleGluGlyThrValGlu 325
|
|
|
Db 300 AGTATAGATAAATTTTCCACAGCTGGATATAAACCCTGAATCCATAGAAAGGAACTGTGAA 359
|
|
|
QY 326 PheIleAsnValSerPheAsnTyrProSerArgProSerIleIleLeuIleGlyLeu 345
|
|
|
Db 360 TTTAAATAATGTTTCTTCAATATATCCATCAAGACCATCTATCAAGATTCGAAAGGTCG 419
|
|
|
QY 346 AsnLeuArgIleIleYsSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 365
|
|
|
Db 420 AATCTCAGATTAAGTCTGGAGACAGTCGCTTGGTCGGTCTCAATGCGAGTGGGAAG 479
|
|
|
QY 366 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 385
|
|
|
Db 480 AGTACGGTAGCTCAGCTTCTGCAGAGGTTATATGATCCGGATGATGCTTTATCATGGTG 539
|
|
|
QY 386 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal 405
|
|
|
Db 540 GATGAGATGACATCAGAGCTTTAAATGTGCGCATTTATCGAGACCATATTTGGAGTGGTT 599
|
|
|
QY 406 SerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAsp 425
|
|
|
Db 600 AGTCAAGAGCTGTTTGTTCGGGACCATCATCAATATCAAGTATGCGAGAT 659
|
|
|
QY 426 AspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIle 445
|
|
|
Db 660 GATGTGCTACTGATGAAGAGATGGAGAGCAGCAAGGGAAGCAATGCGTATGATTTATC 719
|
|
|
QY 446 MetGluPheProAsnLysPheAsnThrLeuValGlyLysGlyAlaGlnMetSerGly 465
|
|
|
Db 720 ATGGAGTTTCTCTAATAA-TTTTAAATACATTTGGTAGGGGAANAAGGAGCTCCAAATGAGTGA 778
|
|
|
QY 466 GlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIle 485
|
|
|
Db 779 AGCAGAACCG-AGGATCCCAATGTCTCTGCTTGTAGTTTGGAAACCCAGAT-CTGAAT 836
|
|
|
QY 486 LeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaLeu 505
|
|
|
Db 837 TTAGATGAAGCTACGTCTGCTCCCTGGATTCAAGAAACCAAGTCAACCTGGTCA--GCTGGGCTG 894
|
|
|
QY 506 GluIleAspThrProArgTyrSerPhe 514
|
|
|
Db 895 GAGAAGGAAACCCCGGNAATTCTTT 921
|
|
|
RESULT 3
BM016204 760 bp mRNA linear EST 30-OCT-2001
LOCUS 603642659F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:5418615 5',
DEFINITION mRNA sequence.
ACCESSION BM016204
VERSION BM016204.1 GI:16530558
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 760)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue procurement: DCTD/DTF
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LIAW12068 row: e column: 16
High quality sequence stop: 740.
Location/Qualifiers
```

FEATURES

```
source 1..760
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NIH_MGC_87"
/tissue_type="mammary adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: breast; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.383 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT 229 a 150 c 187 g 193 t
ORIGIN
Alignment Scores:
Pred. No.: 7,18e-112 Length: 760
Score: 1029.00 Matches: 209
Percent Similarity: 98.12% Conservatives: 0
Best Local Similarity: 98.12% Mismatches: 2
Query Match: 39.70% Indels: 2
DB: 13 Gaps: 0
US-09-873-409-8 (1-514) x BM016204 (1-760)
QY 303 LysIleProSerIleAspAsnPheSerThrAlaGlyTyrIleProGluSerIleGluGly 322
|
|
|
Db 3 AGAAACCCAGTATAGATAAATTTTCCACAGCTGGATATAAACCCTGAATCCATAGAAAGGA 62
|
|
|
QY 323 ThrValGluPheIleAsnValSerPheAsnTyrProSerArgProSerIleIleLeu 342
|
|
|
Db 63 ACTGTGGAATTTTAAATAATGTTTCTTCAATATATCCATCAAGACCATCTATCAAGATTCG 122
|
|
|
QY 343 LysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGly 362
|
|
|
Db 123 AAAGTCTGATCTCAGATTTAAGTCTGGAGACAGTGCCTTGGTTCGCTCAATGGC 182
|
|
|
QY 363 SerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPhe 382
|
|
|
Db 183 AGTGGGAAGAGTAGCTAGTCCAGCTTCTCGAGAGCTTATATGATCCGGATGATGCTTT 242
|
|
|
QY 383 IleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIle 402
|
|
|
Db 243 ATCATGCTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCAATTTATCGAGACCATATN 302
|
|
|
QY 403 GlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyr 422
|
|
|
Db 303 GGAGTGTGTTAGTCAAGAGCTGTTTGTTCGGGACCATCATAGTAACTATATCAAGTAT 362
|
|
|
QY 423 GlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyr 442
|
|
|
Db 363 GGACGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAAGGGAAGCAATGCGTAT 422
|
|
|
QY 443 AspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyLysGlyAlaGln 462
|
|
|
Db 423 GATTTTATCATGAGTTCCTTAATAATTTTAAATACATTTGGTAGGGGAAAAGAGCTCAA 482
|
|
|
QY 463 MetSerGlyGlyGlnLysGlnArgIleAlaIle-AlaArgAlaLeuValArgAsnProLys 482
|
|
|
Db 483 ATGAGTGGAGGGCAGAAACAGAGATCGCAATTTGGCTTGGCTTGTTCGAAACCCCAA 542
|
|
|
QY 482 sIleLeuIleLeu-AspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValG 502
|
|
|
Db 543 GATTTCTGATTTTCAGATGAGGCTACGCTCGCTCGCTGATTCAGAAAGCAAGTCAGTGTTC 602
|
|
|
QY 502 lnAlaAlaLeuGluLysAspThrProArgTyrSer 513
|
|
|
Db 603 AAGCTGCACTGGAGAAGGATACCCCGAGGTATTTC 637
|
|
|
RESULT 4
BQ882401 929 bp mRNA linear EST 16-AUG-2002
LOCUS BQ882401
DEFINITION AGENCOURT_8627902 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:6291782
```


Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 948)

NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC/DCFT/DTP
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: LLAM12233 row: b column: 14
 High quality sequence stop: 569.
 Location/Qualifiers
 1..948
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5539117"
 /clone_lib="NIH_MGC_72"
 /tissue_type="melanotic melanoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: skin; Vector: pCMV-SPORT6; Site: 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT Average insert size 2 kb. Library constructed by Life Technologies."

BASE COUNT	285 a	187 c	194 g	282 t
ORIGIN				
Alignment Scores:				
Pred. No.:	3,98e-91	Length:	948	
Score:	856.50	Matches:	186	
Percent Similarity:	89.67%	Conservative:	5	
Best Local Similarity:	87.32%	Mismatches:	14	
Query Match:	33.04%	Indels:	8	
DB:	13	Gaps:	1	

US-09-873-409-8 (1-514) x BM471690 (1-948)

QY 266 ValPheSerValIleHisSerTyrCysIleGlyAlaAlaValProHisPheGlu 285
 Db 205 GTTTTCTTTAGTGTAAATCCATAGCAGTATTGTCATTGGAGCAGCAGTCCCTCCTCTTGA 264
 QY 286 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysPro 305
 Db 265 ACCTTCGCAATAGCCGAGGAGTCGCTTTTCATATTTCCAGGTATTGATGAAGAACCC 324
 QY 306 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 325
 Db 325 AGTATAGATAACTTTTCCACAGCTGGATATAAACCCTGAATCCATAGAGAACTGTGGAA 384
 QY 326 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu 345
 Db 385 TTTTAAAAATGTTCTTTTCAATTATCCATCAAGACCATCTATCAAGATTCTGAAGAGTCTG 444
 QY 346 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 365
 Db 445 AATCTCAGAATTAAGTCTCGAGAGACAGTCGCTTGGTGGCTCTCAATGGCAGTGGGAAG 504
 QY 366 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 385
 Db 505 AGTACGGTAGTCCAGCTTCTCGAGAGGTTATATGATCCGAGTATGGCTTTTATCATGGTG 564
 QY 386 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIle-GlyValVa 405
 Db 565 GATGAGAATGACATCAGAGCTTTAAATGTGCGCATTTATCGAGACCATATTTGGAGTGT 624
 QY 405 lSerGlnGluPro--ValLeuPheGlyThrThrIleSerAsnIleLysTyr-GlyAr 424
 Db 625 TAGTCAAGAGCCCTGGTTTGTTCGGGACACCATCAGTAACAATATCAATATGAGTATGGAGC 684

FEATURES
 source
 1..545
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="DKFZp434C1815"
 /clone_lib="434 (synonym: htes3)"
 /tissue_type="testis"
 /dev_stage="adult"
 /lab_host="DH10B"
 /notes="Vector: pSport1; Site_1: NotI; Site_2: SalI"

BASE COUNT 171 a 89 c 138 g 147 t
 ORIGIN

Alignment Scores:
 Pred. No.: 7.34e-99 Length: 545
 Score: 918.00 Matches: 181
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 35.42% Indels: 0
 DB: 9 Gaps: 0

US-09-873-409-8 (1-514) x AL040762 (1-545)

QY 306 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 325
 Db 3 AGTATAGATAACTTTCCACAGCTGGATATAAACCCTGAATCCATAGAGAACTGTGGAA 62
 QY 326 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu 345
 Db 63 TTTAAAAATGTTCTTTTCAATTCATCAAGACCATCTATCAAGATTCTGAAGAGTCTG 122
 QY 346 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 365
 Db 123 AATCTCAGAATTAAGTCTGGAGAGACAGTCGCTTGGTCTCAATGGCAGTGGGAAG 182
 QY 366 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 385
 Db 183 AGTACGGTAGTCCAGCTTCTGACAGAGGTATATGATCCGATGATGCTTTATCATGGTG 242
 QY 386 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal 405
 Db 243 GATGAGAATGACATCAGAGCTTTAAATGTGCGGATATTCGAGACCATATTTGGAGTGGT 302
 QY 406 SerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAsp 425
 Db 303 AGTCAAGAGCTGTTTGTTCGGACACCATCAGTAACAATATCAAGTATGACGACGAT 362
 QY 426 AspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIle 445
 Db 363 GATGTGACTGATGAAGAGATGAGAGAGCAGCAGAGGGAAGCAATGCGTATGATTTTATC 422
 QY 446 MetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGly 465
 Db 423 ATGGAGTTTCCCTAAATAAATTTAATACATGTTGGTGGGAAAAGGAGCTCAAAATGAGTGA 482
 QY 466 GlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIle 485
 Db 483 GGGCAGAAACAGAGATCGCAATTGCTCGTGCTTAGTTTGAATCCCAAGATTCTGATT 542
 QY 486 Leu 486
 Db 543 TTA 545

RESULT 6
 BM471690
 LOCUS
 DEFINITION AGENCOURT 6465349 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5539117
 5', mRNA sequence.
 ACCESSION BM471690
 VERSION BM471690.1 GI:18520732
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens


```

Qy 424 gAspValThrAspGluGlu-MetGluArgAlaAlaArgGluAla-AsnAlaTyrAsp 443
Db 685 AGATGATGGGACTGATGAACAAATGGAAAAGCAACAGGGAACCAAGCGCAATGAA 744
Qy 444 PheIleMetGlu-PheProAsnLysPheAsnThrLeuVal---GlyGluLysGlyAlaG1 462
Db 745 TTTTATCATGGAATTTCCCAAAAAAATTAATACATTGGAAGGGGAAAAAAGGACCTTCA 804
Qy 462 nMetSerGlyGlyGlnLysGlnArg 470
Db 805 AGGAATCGAAGGCGCAAAACCAAG 829

RESULT 7
LOCUS BF584668 894 bp mRNA linear EST 12-DEC-2000
DEFINITION 602098406F1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4218385 5',
mRNA sequence.
ACCESSION BF584668
VERSION BF584668.1 GI:11658386
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-x@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM9798 row: 1 column: 02
High quality sequence stop: 651.

FEATURES
source
1..894
location/Qualifiers
/organism="Mus musculus"
/strain="FVB/N"
/db xref="taxon:10090"
/clone="IMAGE:4218385"
/clone_lib="NCI_CGAP_Co24"
/lab host="DH10B (TI phage-resistant)"
/notes="Organ: colon; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.6 kb. Constructed by Life
Technologies. Note: this is a NCI CGAP Library."
BASE COUNT 263 a 203 c 242 g 185 t
ORIGIN
1
Alignment Scores:
Pred. No.: 1,85e-78 Length: 894
Score: 750.00 Matches: 156
Percent Similarity: 84.94% Conservative: 47
Best Local Similarity: 65.27% Mismatches: 36
Query Match: 28.94% Indels: 4
DB: 12 Gaps: 0

US-09-873-409-8 (1-514) x BF584668 (1-894)

Qy 266 ValPheSerValIleHisSerTyrCysTyrGlyAlaAlaValProHisPheGlu 285
Db 5 GTCTCTTTCCGTTTAAT-GGAGCATTCAGTGTGACAGGCACTCTCCAAATATTGAA 63
Qy 286 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAapLysLysPro 305
Db 64 GCCTTCGCCAATGCACGAGGAGCAGCTTATGAAGTCTTCAAAATATTGATAAGGCC 123

```

```

Qy 306 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 325
Db 124 AGTATAGACAGCTTCTCAAAGAGTGGGCAAAACACCAAGCAATACAGAAATCTGGAA 183
Qy 326 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLysGlyLeu 345
Db 184 TTTTAAAGATATTCACTTCAGTTACCCATCTCGAAAAAGAGTTCCAGATCTTTGAAGGCCCTC 243
Qy 346 AspLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 365
Db 244 AATCTGAGGTGAAGAGCGGACAGCGGTGGCCCTGGTGGCAACAGTGGCTGTGAAAA 303
Qy 366 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 385
Db 304 AGCACAACGTCTCAGCTGATGCAAGGCTCTACGACCCCTAGATGGCTGTCAGTATC 363
Qy 386 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal 405
Db 364 GACGGACAGGACATCAGAACCATCAATGTGAGGTATCTGAGGGAGATCATTTGTGTGGTG 423
Qy 406 SerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAsp 425
Db 424 AGTCAGAACCTGTGTGTTGCCACACGATCGCGAGACATTCGCTATGGCCGAGAA 483
Qy 426 AspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIle 445
Db 484 GATGTCAACATGGATGAGATTGAGAAAGCTGTCAAGAGCAATGCTATGACTTTCATC 543
Qy 446 MetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGly 465
Db 544 ATGAAATCGCCCAACCAATTTGACCTCTGGTGGTGAGAGGGCGCGCAGCTGAGTGGG 603
Qy 466 GlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIle 485
Db 604 GGAACAGAAACAGAGATCGC-ATTGCCGGGCCCTGTGTCGCAATCCCAAGTGCCTTTC- 661
Qy 486 LeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 504
Db 662 GTGACGAGACCCACCTCAGC-CTGGATACAGAAAGTGAAGCTGGGTTCAGGCCGCA 717

RESULT 8
LOCUS AK020318 872 bp mRNA linear HTC 19-JAN-2002
DEFINITION Mus musculus adult male epididymis cDNA, RIKEN full-length enriched
library, clone:9230106F14:similar to MULTIDRUG RESISTANCE PROTEIN 3
(P-GLYCOPROTEIN 3) (MDR1A), full insert sequence.
ACCESSION AK020318
VERSION AK020318.1 GI:12860872
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (strain:C57BL/6J) adult male epididymis cDNA to mRNA,
clone lib:RIKEN full-length enriched mouse cDNA library
clone:9230106F14.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,

```


Yanamoto, R., Matsumoto, H., Sakeguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

RIXEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer

Genome Res. 10 (11), 1757-1771 (2000)

20530913

11076851

4

Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, I., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, S., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, H., Fleischmann, W., Gaasterland, T., Giesi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsumo, Y., Nikaïdo, I., Pesole, G., Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whitaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.

TITLE	JOURNAL	MEDLINE	PUBMED	REFERENCE	AUTHORS
1. <i>Journal of the American Medical Association</i>					
2. <i>British Medical Journal</i>					
3. <i>New England Journal of Medicine</i>					
4. <i>Lancet</i>					
5. <i>Annals of Internal Medicine</i>					
6. <i>Journal of Clinical Investigation</i>					
7. <i>Journal of Biological Chemistry</i>					
8. <i>Journal of the Neurological Sciences</i>					
9. <i>Journal of the American Society of Nephrology</i>					
10. <i>Journal of the American Society of Hypertension</i>					
11. <i>Journal of the American Society of Gerontology</i>					
12. <i>Journal of the American Society of Geriatrics</i>					
13. <i>Journal of the American Society of Geriatrics</i>					
14. <i>Journal of the American Society of Geriatrics</i>					
15. <i>Journal of the American Society of Geriatrics</i>					
16. <i>Journal of the American Society of Geriatrics</i>					
17. <i>Journal of the American Society of Geriatrics</i>					
18. <i>Journal of the American Society of Geriatrics</i>					
19. <i>Journal of the American Society of Geriatrics</i>					
20. <i>Journal of the American Society of Geriatrics</i>					
21. <i>Journal of the American Society of Geriatrics</i>					
22. <i>Journal of the American Society of Geriatrics</i>					
23. <i>Journal of the American Society of Geriatrics</i>					
24. <i>Journal of the American Society of Geriatrics</i>					
25. <i>Journal of the American Society of Geriatrics</i>					
26. <i>Journal of the American Society of Geriatrics</i>					
27. <i>Journal of the American Society of Geriatrics</i>					
28. <i>Journal of the American Society of Geriatrics</i>					
29. <i>Journal of the American Society of Geriatrics</i>					
30. <i>Journal of the American Society of Geriatrics</i>					
31. <i>Journal of the American Society of Geriatrics</i>					
32. <i>Journal of the American Society of Geriatrics</i>					
33. <i>Journal of the American Society of Geriatrics</i>					
34. <i>Journal of the American Society of Geriatrics</i>					
35. <i>Journal of the American Society of Geriatrics</i>					
36. <i>Journal of the American Society of Geriatrics</i>					
37. <i>Journal of the American Society of Geriatrics</i>					
38. <i>Journal of the American Society of Geriatrics</i>					
39. <i>Journal of the American Society of Geriatrics</i>					
40. <i>Journal of the American Society of Geriatrics</i>					
41. <i>Journal of the American Society of Geriatrics</i>					
42. <i>Journal of the American Society of Geriatrics</i>					
43. <i>Journal of the American Society of Geriatrics</i>					
44. <i>Journal of the American Society of Geriatrics</i>					
45. <i>Journal of the American Society of Geriatrics</i>					
46. <i>Journal of the American Society of Geriatrics</i>					
47. <i>Journal of the American Society of Geriatrics</i>					
48. <i>Journal of the American Society of Geriatrics</i>					
49. <i>Journal of the American Society of Geriatrics</i>					
50. <i>Journal of the American Society of Geriatrics</i>					
51. <i>Journal of the American Society of Geriatrics</i>					
52. <i>Journal of the American Society of Geriatrics</i>					
53. <i>Journal of the American Society of Geriatrics</i>					
54. <i>Journal of the American Society of Geriatrics</i>					
55. <i>Journal of the American Society of Geriatrics</i>					
56. <i>Journal of the American Society of Geriatrics</i>					
57. <i>Journal of the American Society of Geriatrics</i>					
58. <i>Journal of the American Society of Geriatrics</i>					
59. <i>Journal of the American Society of Geriatrics</i>					
60. <i>Journal of the American Society of Geriatrics</i>					
61. <i>Journal of the American Society of Geriatrics</i>					
62. <i>Journal of the American Society of Geriatrics</i>					
63. <i>Journal of the American Society of Geriatrics</i>					
64. <i>Journal of the American Society of Geriatrics</i>					
65. <i>Journal of the American Society of Geriatrics</i>					
66. <i>Journal of the American Society of Geriatrics</i>					
67. <i>Journal of the American Society of Geriatrics</i>					
68. <i>Journal of the American Society of Geriatrics</i>					
69. <i>Journal of the American Society of Geriatrics</i>					
70. <i>Journal of the American Society of Geriatrics</i>					
71. <i>Journal of the American Society of Geriatrics</i>					
72. <i>Journal of the American Society of Geriatrics</i>					
73. <i>Journal of the American Society of Geriatrics</i>					
74. <i>Journal of the American Society of Geriatrics</i>					
75. <i>Journal of the American Society of Geriatrics</i>					
76. <i>Journal of the</i>					

TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409 (6821), 685-690 (2001)

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

11217631
5 (bases 1 to 872)

Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Anoh, H., Arai, A.,
Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bulc, C.,
Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanaagaki, T.,
Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiraoka, T., Hori, F.,
Hume, D., Imotani, K., Ishii, Y., Itoch, M., Izawa, M., Kasukawa, T.,
Kato, H., Kawaj, J., Kojima, Y., Konno, H., Kouda, M., Koya, S.,
Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K.,
Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J.,
Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,
Schriml, L., Shibata, K., Shibata, Y., Shingawa, A., Shiraki, T.,
Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F.,
Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamanaka, I.,
Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and
Havashizaki, Y.

направил, т.
Direct Submission

Submitted (18-AUG-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9222

COMMENT

CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5', GAGAGAGAGAAGTCCAGACGCTCTTTTTTTTTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 185.2. Second strand cDNA was prepared with the primer adapter of sequence [5', GAGAGAGAGATTCGAGCTTAATTAATTAATCCCCCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified pluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI. Host: DH10B.

```

FEATURES             Location/Qualifiers
     .1..872
     /organism="Mus musculus"
     /strain="C57BL/6J"
     /db_xref="PANTOM.DB:9230106F14"
     /db_xref="MGI:1911565"
     /db_xref="taxon:10090"
     /clone="9230106F14"
     /sex="male"
     /tissue type="epididymis"
     /clone_lib="RIKEN full-length enriched mouse cDNA library"
     86..>872
     /dev_stage="adult"
     /notes="data source:SPTR, source key:P21447, evidence:ISS
     putative
     similar to MULTIDRUG RESISTANCE PROTEIN 3 (P-GLYCOPROTEIN
     3) (MDR1A)"
     /codon_start=1
     /protein_id="BAB32070.1"
     /db_xref="GI:12860873"
     /db_xref="MGI:1924956"
     /translation="MANSRTNGSLQETNQRYGLQVQPKVQNAQVGPPIEIRPADNL
     DIVLMTGLIASMGVPLMSLVGLSGDHLINGCLVQTNRTKTCNQSCQTEKLN
     DIIVLTLYIGIGAAALIFGYVOISFWVITAAQRTIRKQPFHSLAQDISWFDGSD
     ICELNRTMGDINKCLGDKIPLMFQNIISGFSIGLVLSLKSLKSLVLVLTSPPLI
     MASSALTRGDIINSLSKELDAYSKAGAVAEAUSSITQTVAFGA"
262 a 178 c 187 g 245 t

```

Alignment Scores:

Pred. No.:	8,068-78	Length:	872
Score:	744.50	Matches:	151
Percent Similarity:	80.3%	Conservative:	21
Best Local Similarity:	70.5%	Mismatches:	27
Query Match:	28.7%	Indels:	15
DB:	11	Gaps:	2

US-09-873-409-8 (1-514) x AK020318 (1-872)

[illegible][illegible]

Db	710	CTGGTGTTCTGTCTACATCTCTCTCATATGCTTTCATCGGACGTGCTCTAGGATG	769	JOURNAL MEDLINE PUBMED REFERENCE AUTHORS	Nature 409 (6821), 685-690 (2001) 21085660 11217851 5 (bases 1 to 2676) Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A., Arakawa,T., Baldarelli,R., Bono,H., Brownstein,M., Bult,C., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hangaki,T., Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hiraoka,T., Hori,F., Hume,D., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numata,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J., Saito,H., Saito,R., Sakai,K., Sano,H., Sasaki,D., Schriml,L., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yamanaka,I., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y. Direct Submission Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gscc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216) Please visit our web site (http://genome.gsc.riken.go.jp/) for further details. CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'-GAGGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5'-GAGGAGAGATTCGAGTTAAATTAATTAATCCCTCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. cDNA of size comprised between 3 and 7 kb was selected before cloning. Vector: a modified pBluescript KS(+) after bulk excision from lambdaBda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI. Host: DRI0B. Location/Qualifiers 1. .2676 /organism="Mus musculus" /strain="C57BL/6J" /db_xref="FANTOM DB:3222401P09" /db_xref="MGI:1907488" /db_xref="taxon:10090" /clone="3222401P09" /tissue type="head" /dev_stage="14, 17 days embryo"
Qy	166	ValleSerLeuThrSerLySGLuSerAlaTySerLySAlaGlyAlaValAlaGlu	185		
Db	770	ATTATCTCATTCACCAAGAGCTGGATGCTTATTCACAGCTGGGCTGTGGCTGAA	829		
Qy	186	GluValLeuSerSerIleArgThrValIleAlaPheArgAla	199		
Db	830	GAAGCCTTGTCATCAATCAACACAGTCACAGCCTTTGGAGCC	871		
RESULT 9					
LOCUS	AK014319				
DEFINITION	Mus musculus 14, 17 days embryo head cDNA, RIKEN full-length enriched library, clone:3222401P09:ATP-binding cassette, sub-family B (MDR/TAP), member 8, full insert sequence.				
ACCESSION	AK014319				
VERSION	AK014319.1	GI:12852089			
KEYWORDS	HTC; CAP trapper.				
SOURCE	Mus musculus (strain:C57BL/6J) 14, 17 days embryo head cDNA to mRNA, clone lib:RIKEN full-length enriched mouse cDNA library				
ORGANISM	Mus musculus				
REFERENCE	1	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
AUTHORS	Carninci,P. and Hayashizaki,Y.				
TITLE	High-efficiency full-length cDNA cloning				
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)				
MEDLINE	99279253				
PUBMED	10349636				
REFERENCE	2				
AUTHORS	Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.				
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)				
MEDLINE	20499374				
PUBMED	11042159				
REFERENCE	3				
AUTHORS	Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohata,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsunura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.				
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer				
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)				
MEDLINE	20530913				
PUBMED	11076861				
REFERENCE	4				
AUTHORS	Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y., Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S., Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I., Saito,T., Okazaki,Y., Gotohori,T., Bono,H., Kasukawa,T., Saito,R., Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T., Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H., Kuehl,P., Lewis,S., Matsuo,Y., Nikaudo,I., Pesole,G., Quackenbush,J., Schriml,L.M., Staabli,F., Suzuki,R., Tomita,M., Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H., Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N., Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C., Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D., Hofmann,P., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,S., Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P., Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H., Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H., Toyooka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L., Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawai,J.H., Kohsaki,S. and Hayashizaki,Y. Functional annotation of a full-length mouse cDNA collection				
TITLE					
FEATURES					
source					
gene					
CDS					
member 8					
data source:MGI					
putative					
/codon_start=1					
/protein_id="BAB29270.1"					
/db_xref="GI:12852090"					
/db_xref="MGI:1351667"					
/translation="MLVHLFRFGIRGGVPVGSQSLRPFQTSAAKSSDRSLSSHLRLTYAQRQLRAHLPRAPPASHWSPSCWVGTVLPVAVLWHPRLCLALCEAKESPQAQTRAPLRNFNKLFWHLPHLLALGAALVQVPLLLGQVLEIVAKYTRDHMGFSVSSKLVOLLVYQVGLLTGYLVLLSHIGERWMDNRKALFSSLLRODIAFFDAKKTGQLVSRLLTVDQBFKSSFKLVISQGLRSCVTGSLVSLWLSPLRLTLMVAVPALMGVGTLMGSLGRLKLRSCQEQIARATGVADALGNVTRVAFAMEKREERYQAELESCCCAEELGQVGLFQGLSNATFNQVLTGLTFIGSSLVAGQQLKGGDLMSFLVASQTQVRSMAISLVLFQVVRGLSAGARVFYFNALSFPVPLTGTGCTLPNKDI					

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
 Email: c9apbs-f@mail.nih.gov
 Tissue Procurement: The Cepko Laboratory
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov
 Plate: LLAM10394 row: b column: 22
 High quality sequence stop: 698.

FEATURES
 source
 1..803
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="IMAGE:4511157"
 /clone_lib="NIH MGC 94"
 /tissue_type="retina"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: eye; Vector: pCMV-SPORT6; Site: 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 3.3 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH MGC Library."
 240 a 162 c 205 g 196 t

BASE COUNT 240 a 162 c 205 g 196 t
ORIGIN

Alignment Scores:
 Pred. No.: 3,636-69 Length: 803
 Score: 671.50 Matches: 149
 Percent Similarity: 76.15% Conservative: 49
 Best Local Similarity: 57.31% Mismatches: 55
 Query Match: 25.91% Indels: 10
 DB: 12 Gaps: 2

US-09-873-409-8 (1-514) x BG298756 (1-803)

QY 168 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluVal 187
 DB 15 TCATTTACTGATGAAGAACTCCATGCTTATGCAAAAGCTGAGCAGTGTCTGAAGAATC 74
 QY 188 LeuSerSerLeuArgThrValLeuAlaPheArgAlaGlnGluLysGluLeuGlnArgTyr 207
 DB 75 TTAGCAGCCATCAGAACTGTGATTCCTTGA-GGACAAAGAGGAACTTGAAGGTAC 133
 QY 208 ThrGlnAsnLeuLysAspAlaLysAspPheGlyLeuLysArgThrLeuAlaSerLysVal 227
 DB 134 AATAACAACCTTGAAGAAGCTAAAGGCTGGGATAAAGAAAGTATCACGGCCAAATC 193
 QY 228 SerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrGly 247
 DB 194 TCCATGGTGCAGCTTTCTCTTATCTATGTCATCATATGCTCTGGCATCTCTGGTATGG 253
 QY 248 ThrSerLeuLeuAsnGlyGluProGlyTyrThrLeuGlyThrValLeuAlaValPhe 267
 DB 254 ACTTCCTGGTCATCTCAAGAA-----TACTCTAT-GGACAAAGTCTCACTGCTTC 306
 QY 268 PheSerValLeuHisSerSerTyrCysAlaGlyAlaValProHisPheGluThrPhe 287
 DB 307 TTTTCGCTGTTAATGGAGCATTCAGTGTGGACAGGCATCTCCAAATATTGAAGCCTTC 366
 QY 288 AlaLeuAlaArgGlyAlaAlaPheHisLeuPheGlnValLeuLysLysProSerile 307
 DB 367 GCCAATCCAGCAGGAGCAGCTTATGAGTCTTCAAAATAATATGATAAGCCAGTATA 426
 QY 308 AspAsnPheSerThrAlaGlyTyrLysProGluSerLeuGluGlyThrValGluPheLys 327
 DB 427 GACAGCTTCTCAAGAGTGGCCAAACCAACACACATACAAAGAAATCTGGAATTTAAG 486
 QY 328 AsnValSerPheAsnTyrProSerArgProSerLeuLysLysLeuLysGlyLeuAsnLeu 347
 DB 328 AsnValSerPheAsnTyrProSerArgProSerLeuLysLysLeuLysGlyLeuAsnLeu 347

DB 487 AATATTCACTTCAGTTTACCATCTCGAAAGAAGTTTCAGATCTTGAAGGCCCTCAATCTG 546
 QY 348 ArgLeuLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThr 367
 DB 547 AGGTGAAGAGCGGACAGACGGTGGCCCTGGTTGGCAACAGATGGCTGTGAAAGACACA 606
 QY 368 ValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheLeuMetValAspGlu 387
 DB 607 ACTGTCCAGCTGATCAAGAGCTCTACGACCCCTA-GATGGCATGCTCAGTATCGACGGA 665
 QY 388 AsnAspIleArgAlaLeu-LeuAsnValArgHis-TyrArg-AspHisIleGlyValValSer 406
 DB 666 CAGGACATCAGAACCACTCAATGTCAGGTATCTGGAGGGGATCAT---GGGGTGGTCA 722
 QY 407 Gln-GluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 423
 DB 723 CAAGGAACCTGGCTTGTGTTGCCAACCCAGATCGCGAGAAATTCGTATATGGG 774

RESULT 11
 AL520322
 LOCUS AL520322 LTI_NFL004_NBC2 Homo sapiens cDNA clone CS0DB006YC15 5
 DEFINITION prime, mRNA sequence.
 ACCESSION AL520322
 VERSION AL520322.1 GI:12783815
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 943)
 AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
 source
 1..943
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CS0DB006YC15"
 /clone_lib="LTI_NFL004_NBC2"
 /sex="male"
 /tissue_type="neuroblastoma cells"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"
 253 a 225 c 253 g 211 t 1 others

BASE COUNT 253 a 225 c 253 g 211 t
ORIGIN

Alignment Scores:
 Pred. No.: 1,478-67 Length: 943
 Score: 659.00 Matches: 131
 Percent Similarity: 72.36% Conservative: 47
 Best Local Similarity: 53.25% Mismatches: 66
 Query Match: 25.42% Indels: 2
 DB: 9 Gaps: 1

US-09-873-409-8 (1-514) x AL520322 (1-943)

QY 264 LeuAlaValPhePheSerValLeuHisSerSerTyrCysAlaGlyAlaAlaValProHis 283
 DB 2 CTGTATTATTTCAGCTGTGTGTTGTTGTCATCGCGTGGGCAAGTCAGTTTCATTT 61

QY 284 PheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLys 303
 Db 62 GCTCTGACTATGCAAGCCAAATATCAGCAGCCACATCATCATGATCATTTGAAAAA 121
 QY 304 LysProSerIleAspAsnPheserThrAlaGlyTyrLysProGluSerIleGluGlyThr 323
 Db 122 ACCCTTTGATTCAGACAGTACAGCAGGAGCGCTTAATCCGAAACATTTGGAAGGAAT 181
 QY 324 ValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLysLys 343
 Db 182 GTCACATTTGGTGAAGTTGATTCACATATCCACCCGCGGACATCCAGCTGCTTCA 241
 QY 344 GlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySer 363
 Db 242 GCAGTGCAGCTGAGGTGAAGAGGCGCAGACGCTGGCTCTGGTGGCAGCAGTGGCTGT 301
 QY 364 GlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspGlyPheIle 383
 Db 302 GGAAGAGCAGACATGCTGCTGAGGCGGTCTTACGACCCCTTGGCAGGAAAGTG 361
 QY 384 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 403
 Db 362 CTGCTTGATGGCAAGAAATAAAGCAGCTGAATGTTTCAGTGGCTCCGACACACCTGGGC 421
 QY 404 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 423
 Db 422 ATCGTCTCCAGAGGCCATCTGTTTGATCGCAGCATGCTGAGAACATTCCTATGGA 481
 QY 424 ArgAspAsp-----ValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAla 441
 Db 482 GACACAGCGGGTGTGTGCACAGGAAGAGATCGTGAGGGCAGCAAGAGGCGCAACATA 541
 QY 442 TyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAla 461
 Db 542 CATGCTTCATCGAGTCAGCTGCTTAATATATAGACATAAGTAGGAGACAAAGAACT 601
 QY 462 GlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnPro 481
 Db 602 CAGCTCTGCTGGTGGCCAGAAACACCATTTGCCATGCTGCTGCTGTTGTAGACAGCT 661
 QY 482 LysIleLeuLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaVal 501
 Db 662 CATATTTTGGTTTGGTGAAGCCAGCTCAGCTCTGATACAGAAAGTGAAAGGTTGTC 721
 QY 502 GlnAlaAlaLeuGluLys 507
 Db 722 CAAGAAGCCCTGGACAAA 739

RESULT 12
 AA243820 405 bp mRNA linear EST 06-AUG-1997
 LOCUS z67906.r1 Soares NHMPu S1 Homo sapiens cDNA clone IMAGE:668506 5',
 DEFINITION similar to SW:MDRI_MOUSE P06795 MULTIDRUG RESISTANCE PROTEIN 1 ;,
 mRNA sequence.

ACCESSION AA243820
 VERSION AA243820.1 GI:1874631
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 405)
 REFERENCES Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
 Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B.,
 Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie,
 T., Waterston,R. and Wilson,R.
 WashU-Merck EST Project 1997
 TITLE Unpublished (1997)
 JOURNAL Contact: Wilson RK
 COMMENT Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800

FEATURES
 source
 1..405
 /organism="Homo sapiens"
 /db_xref="GDB:5562473"
 /db_xref="taxon:9606"
 /clone="IMAGE:668506"
 /clone_lib="Soares NHMPu S1"
 /tissue_uterus="Pooled human melanocyte, fetal heart, and
 pregnant uterus"
 /lab_host="DH10B"
 /note="Organ: mixed (see below); Vector: pT7T3D-Pac
 (Pharmacia) with a modified polylinker; Site 1: Not I;
 Site 2: Eco RI; Equal amounts of plasmid DNA from three
 normalized libraries (melanocyte 2NBHM, pregnant uterus
 NBHPU, and fetal heart NBHH19W) were mixed, and ss circles
 were made in vitro. Following HAP purification, this DNA
 was used as tracer in a subtractive hybridization
 reaction. The driver was PCR-amplified cDNAs from pools of
 5,000 clones made from the same 3 libraries. The pools
 consisted of I.M.A.G.E. clones 260232-265223,
 340488-345479, and 484488-489479."

BASE COUNT 115 a 81 c 85 g 124 t
 ORIGIN

Alignment Scores:
 Pred. No.: 1.44e-67 Length: 405
 Score: 654.00 Matches: 128
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 99.22% Mismatches: 0
 Query Match: 25.23% Indels: 0
 DB: 9 Gaps: 0

US-09-873-409-8 (1-514) x AA243820 (1-405)
 QY 264 LeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHis 283
 Db 18 CTCTCTGTTTCTTAGTGTAAATCATACAGTATTGTCATTGGACGACAGTCCCTCAC 77
 QY 284 PheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLys 303
 Db 78 TTTGAACCTTCGCAATAGCCCGAGGAGCTGCTTTCATATTTCCAGGTATTGATAAG 137
 QY 304 LysProSerIleAspAsnPheserThrAlaGlyTyrLysProGluSerIleGluGlyThr 323
 Db 138 AAACCCAGTATAGTAACCTTTCCACAGCTGGATATAAACCTGAATCCATAGAAGGAAC 197
 QY 324 ValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLysLys 343
 Db 198 GTGGAAATTTAAATAATGTTCTTTCATATTCATCAAGACCATCTATCAAGATTCGAAA 257
 QY 344 GlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySer 363
 Db 258 GTCTGAAATCTCAGAATTAACTCTGGAGAGACAGTCCCTTGGTCGCTTCAATGSCAGT 317
 QY 364 GlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspGlyPheIle 383
 Db 318 GGAAGAGTACGCTAGTCTGAGCTTCTGAGAGGATATATGATCCGAGTATGATGCTTTATC 377
 QY 384 MetValAspGluAsnAspIleArgAla 392
 Db 378 ATGGTGGTGAAGATGACATCAGAGCT 404

RESULT 13
 BU075336 550 bp mRNA linear EST 11-DEC-2001
 LOCUS BU075336 NIBB Mochii normalized Xenopus tailbud library Xenopus
 DEFINITION laevis cDNA clone XL057d10 5', mRNA sequence.

Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available royalty-free through LLNL ; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 1572 Std Error: 0.00
 Seq primer: -28ml3 rev2 ET from Amersham.
 Location/Qualifiers
 1..405

US-09-873-409-8 (1-514) x BG587938 (1-780)

```
QY 277 IleGlyAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHis 296
Db 3 TTGGGCAATCTGCACGAGCATGGCTGCATTACAAAGCTAGATGGCGCTGTAAG 62
QY 297 IlePheGlnValIleAspLysPheProSerIleAspAsnPheSerThrAlaGlyTyrLys 316
Db 63 ATTTCGGGATAATTGATCACCAGCCTGGTATAGATAGAAACAGTGAATCTGGATTGGAA 122
QY 317 ProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArg 336
Db 123 TTAGACAGAGTACTGAGCTGTGTGAACCTGAAATGTGGACTTCTTATCATCATCAAGA 182
QY 337 ProSerIleLysLysLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAla 356
Db 183 CTGAAGTCTGATCTCATGATTTCTCTGAGTGTCTCCGCGGAAAGACCATAGCT 242
QY 357 LeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyr 376
Db 243 TTAGTTGGTAGCAGCGCTCTGGCAAGAGCAGCTGTGTGTTCTCTTATTGAGAGATTCTAT 302
QY 377 AspProAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArg 396
Db 303 GATCCAACTTCAGGCAAGTAATGTTGGATGGGCATGACATTAATACTTGAAGCTTAAA 362
QY 397 HisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIle 416
Db 363 TGGTTGAGGCAACAATAGACGATGATGAGCCAGAACCTGCTTTGTTGCCACCATGAT 422
QY 417 SerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAla 436
Db 423 CGAGAAATATATCTCTGGGAGCGCTGATGCAACACGAGTTGAGATTGAAGAAGCTGCT 482
QY 437 ArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuVal 456
Db 483 AGGTTTGCTAAATCTCATCTCATCTCATCTCAAACTTCTCGAAGGCTTGAACCTCAGTA 542
QY 457 GlyGlyLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAla 476
Db 543 GGAGAAAGAGGAGTCAACATTTCTGGAGGCAAAAACAAGAATAGCAATAGCAAGGGGA 602
QY 477 LeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGlu 496
Db 603 ATGCTAAAAATCCAGCAATTTCTCTCTCGATGAGGCAACAGTGCATTTGACTCTGAA 662
QY 497 SerLysSerAlaValGlnAlaAlaLeuGluLys 507
Db 663 TCAGAAAAGCTGGTGCAAGAAGCACTTGACCGG 695
```

RESULT 15

BF692596

LOCUS 602248949F1 NIH_MGC_62 Homo sapiens cDNA clone IMAGE:4334100 5',
DEFINITION mRNA sequence.

ACCESSION BF692596

VERSION BF692596.1 GI:11978004

KEYWORDS EST.

SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 559)

NIH-MGC <http://mgs.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs@mail.nih.gov

Tissue Procurement: ATCC/DCTD/DTP

cDNA Library Preparation: CLONETECH Laboratories, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

Plate: LLCMI203 row: a column: 13

High quality sequence stop: 555.

FEATURES
source

1..559
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4334100"
/clone_lib="NIH MGC 62"
/tissue_type="melanotic melanoma, high MDR"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: skin; Vector: pDMR-LIB (Clontech); Site_1:
SfiI (ggcgctcgcc); Site_2: SfiI (ggcgctcgcc);
Double-stranded cDNA was prepared from cell line RNA.
and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CAGCGCATTTATGGCC-3' and 3' adaptor
sequence: 5'-ATTCTAGAGCGGCGGCGGCGCATG-dT(30)BN-3'
(where B = A, C, or G and N = A, C, G, or T). Average
insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA)."

BASE COUNT 161 a 118 c 141 g 139 t

ORIGIN

Alignment Scores:
Pred. No.: 7,08e-64 Length: 559
Score: 625.00 Matches: 134
Percent Similarity: 96.43% Conservatve: 1
Best Local Similarity: 95.71% Mismatches: 4
Query Match: 24.11% Indels: 3
DB: 12 Gaps: 0

US-09-873-409-8 (1-514) x BF692596 (1-559)

```
QY 376 TyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnVal 395
Db 1 TATGATCCGGATGATGGCTTTATCAGGTGGATGAGATGACATCAGAGCTTTAAATGTG 60
QY 396 ArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThr 415
Db 61 CGGCATTTATCGAGACCATAT-GGAGTGGTTAGTCAAGACCTGT-TTGTTCGGGACCAC 118
QY 416 IleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAla 435
Db 119 ATCAGTAAACATATCAAGTATGACGAGATGATGTGACTGATGAAGAGATGGAGAGGCA 178
QY 436 AlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeu 455
Db 179 GCAAGGGAAGCAAAATGCATATGATTTTATCATGGAGTTCCTTAATAAATTAATACATTG 238
QY 456 ValGlyGluLys-GlyAlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAla 475
Db 239 GTAGGGGAAAACAGGAGCTCAAAATGATGAGGCGGCAAGAACAGAGGATCGCAATTCCTCG 298
QY 475 GAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSe 495
Db 299 TGCCTTAGTTCGAACCCCAAGATTTCTGATTTAGATGAGGCTACGCTGCTGCTGATTC 358
QY 495 rGluSerLysSerAlaValGlnAlaLeuGluLysAspThrProArgTyrSerPhe 514
Db 359 AGAAACCGAGTCAGCTGTTCAAGTCGCTGAGAGAGGATACCCCGGAGTATTCATT 416
```

Search completed: March 31, 2003, 13:51:49

Job time : 3547.51 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 31, 2003, 02:24:20 ; Search time 75.2727 Seconds
(without alignments)
2094.145 Million cell updates/sec

Title: US-09-873-409-8
Perfect score: 2592
Sequence: 1 MILGILASLVNGACLPMLP.....SESKSAVOALEKDPTRYSP 514

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DEV=xlh

-MODEL=frame+ p2n.model -DB=xlh
-Q/cgn2_1/USPTO_spool/US09873409/runat_27032003_115422_19259/app_query.fasta_1.7544
-DB-issued Patents NA -Qfmt=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09873409 @CGN 1 1 578 @runat_27032003_115422_19259 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA :
1: /cgn2_6/prodata/2/ina/5A_COMB.seq :
2: /cgn2_6/prodata/2/ina/5B_COMB.seq :
3: /cgn2_6/prodata/2/ina/6A_COMB.seq :
4: /cgn2_6/prodata/2/ina/6B_COMB.seq :
5: /cgn2_6/prodata/2/ina/PCUTS_COMB.seq :
6: /cgn2_6/prodata/2/ina/backfiles1.seq :

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1576	60.8	4233	3	US-09-120-513-1
2	1576	60.8	4233	3	US-09-450-105-1
3	1576	60.8	4669	2	US-08-752-447-1
4	1576	60.8	4669	4	US-09-316-167-1
5	1576	60.6	4646	1	US-08-181-471-2
6	1570	60.6	6505	2	US-08-793-610-5
7	1570	60.6	9318	2	US-08-793-610-6
8	1568.5	60.5	4264	2	US-08-784-649A-1
9	1568.5	60.5	4264	2	US-08-784-649A-5
10	1568	60.5	4669	6	5206352-3
11	1549	59.8	4669	2	US-08-583-276-18
12	975.5	37.6	4224	1	US-08-612-521-1

13	959	37.0	4047	2	US-08-612-734B-1	Sequence 1, Appli
14	939.5	36.2	4800	2	US-08-612-734B-3	Sequence 3, Appli
15	937.5	36.2	4002	2	US-08-996-545-1	Sequence 1, Appli
16	937.5	36.2	4002	2	US-08-996-545-3	Sequence 3, Appli
17	937.5	36.2	4002	4	US-09-328-320-1	Sequence 1, Appli
18	937.5	36.2	4002	4	US-09-328-320-3	Sequence 3, Appli
19	920	35.5	3924	1	US-08-395-246C-1	Sequence 1, Appli
20	902.5	34.8	2726	1	US-08-461-823-1	Sequence 1, Appli
21	859	33.1	6143	1	US-08-612-521-3	Sequence 3, Appli
22	809	31.2	3792	4	US-09-351-224E-10	Sequence 10, Appli
23	770	29.7	3924	2	US-08-996-644-3	Sequence 3, Appli
24	770	29.7	3924	2	US-09-352-552-3	Sequence 3, Appli
25	770	29.7	3927	2	US-08-996-644-1	Sequence 1, Appli
26	770	29.7	3927	3	US-09-352-552-1	Sequence 1, Appli
27	712	27.5	3999	4	US-09-351-224E-9	Sequence 9, Appli
28	669.5	25.8	2376	1	US-08-394-880B-1	Sequence 1, Appli
29	641	24.7	2061	4	US-09-061-764A-17	Sequence 17, Appli
30	639.5	24.7	3909	1	US-08-232-537-1	Sequence 1, Appli
31	635	23.5	1959	4	US-09-061-764A-4	Sequence 4, Appli
32	598	23.1	2244	4	US-09-061-764A-18	Sequence 18, Appli
33	581	22.4	1749	4	US-09-134-001C-1893	Sequence 1893, Ap
34	543.5	21.0	1743	4	US-09-134-001C-2774	Sequence 2774, Ap
35	529	20.4	7760	4	US-08-961-527-63	Sequence 63, Appli
36	521	20.1	5120	3	US-08-772-270A-6	Sequence 6, Appli
37	521	20.1	8370	2	US-08-488-706-1	Sequence 1, Appli
38	498.5	19.2	13188	4	US-08-961-527-70	Sequence 70, Appli
39	498.5	19.2	441529	4	US-09-103-840A-1	Sequence 1, Appli
40	497.5	19.2	7721	3	US-08-772-270A-14	Sequence 14, Appli
41	496.5	19.2	4403765	4	US-09-103-840A-2	Sequence 2, Appli
C 42	495.5	19.1	28804	2	US-08-592-874-1	Sequence 1, Appli
C 43	495.5	19.1	28804	3	US-09-096-942-2	Sequence 2, Appli
C 44	495.5	19.1	28804	3	US-09-096-867-2	Sequence 2, Appli
45	472.5	18.2	2407	2	US-08-895-522-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-120-513-1
; Sequence 1, Application US/09120513
; Patent No. 6025160
; GENERAL INFORMATION:
; APPLICANT: Brun, Kimberly
; APPLICANT: Chenery, Richard
; APPLICANT: Ellens, Harma
; APPLICANT: Field, John
; APPLICANT: Yue, Lin
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE
; TITLE OF INVENTION: SEQUENCES ENCODING RAT MDR1B2 AND
; TITLE OF INVENTION: SCREENING METHODS THEREOF
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY:
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/120,513
; FILING DATE: 22-JUL-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: King, William T
; REGISTRATION NUMBER: 30,954
; REFERENCE/DOCKET NUMBER: GPS0008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5015

TELEFAX: 610-270-5090

TELEX:

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 4233 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: Genomic DNA

US-09-120-513-1

Alignment Scores:

Pred. No.: 2,88e-179 Length: 4233
Score: 1576.00 Matches: 308
Percent Similarity: 75.90% Conserved: 92
Best Local Similarity: 58.44% Mismatches: 105
Query Match: 60.80% Indels: 22
DB: 3 Gaps: 4

US-09-873-409-8 (1-514) x US-09-120-513-1 (1-4233)

QY 1 MetileuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
DB 170 ATGGCTCTGGAACTCTCGCTGCTATCATCCACGAAACCCCTGCTCCCTCCTGATGCTG 229
QY 21 ValLeuGlyGluMetSerAspAsnLeuIle----- 30
DB 230 GTGTTCGGATACATGACAGATAGTTTACCAAGCAGAGACCCCATCTGCCGAGGCTT 289
QY 31 -----SerGlyCysLeuValGlnThrAsnThrTyrSerPhePheArgLeu----- 45
DB 290 ACTAATCAAGTGAATCAACAGATACACAGACCGCTCAGCGACAGCAGCTTGGAGGAGAC 349
QY 46 -----ThrLeuTyrTyrValGlyIleGlyValAlaLeuIlePheGlyTyr 61
DB 350 ATGCCCATGTACCGCTACTATTACCGGCATCGTCCGCTGCTCATCGTTGCCCTAC 409
QY 62 IleGlnIleSerLeuTyrIleAlaAlaArgGlnThrLysArgIleArgLysGln 81
DB 410 ATCCAGGTTTCACTTGGTCTGGCAGCTGGGAGACAAATACACAGATAGGACAGAG 469
QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
DB 470 TTTTTCATGCTCATCAATCAGAGATAGCTGGTTGACGTGAATGACGCTGGGGAG 529
QY 102 LeuAsnThrArgMetThr---AspIleAspIleSerAspGlyIleGlyAspLysIle 120
DB 530 CTCACACCCCGGCTCACAGATGACGTCTCCAAATTAATGACGGAATGGTGACAAACTT 589
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
DB 590 GGAATGTTCTTTCAGTCCATACACACATTTACCCCGGTTTATAATAGATTTATAGT 649
QY 141 GlyTrpLysLeuThrLeuValThrSerThrSerProLeuIleMetAlaSerAlaAla 160
DB 650 GGTGTGAAGCTAAACCTTGTAATTTGGCCGTCAGCCCTCTTATTGGGTGTCATCTGCC 709
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
DB 710 ATGTGGCAAGAGTACTGACTTCATTTACTAATAAGAACTCCAGGCTTATGCGAAAGCT 769
QY 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
DB 770 GGACAGTGTCCGAGAGAGCTTACAGCCATCAGACTGTGATGGTTGGAGGACAA 829
QY 201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyLys 220
DB 830 AAGAAGAACTTGAAGGTACAATAAATTTAGAAGAGCTAAAGAGTTGGCATAAG 889
QY 221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
DB 890 AAGCCATCAGCGCAACATTTCCATAGGTATTGCCCTACCTGTTGGTCTATGCGTCTTAT 949

QY 241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
DB 950 GCATGSCATTTCTGGTATGGACCTCTTGGTCTCTCAATGAA-----TATCTATT 1003
QY 261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
DB 1004 GGACAGTGTCTTACCCTCTTCTCTATTATTATGGGACTTTTCAGTATTGGACATTTA 1063
QY 281 ValProHisPheGluThrPheAlaIleAlaArgGlyValAlaPheHisIlePheGlnVal 300
DB 1064 GCCCAAAACATAGAAGCTTTTGAATGCAAGAGGGGACGCCCTATGAAATCTTCAAGATA 1123
QY 301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
DB 1124 ATTGATAATGAGCCAGCATTTGACAGCTTCTCAACCAAGGGACACAAACACAGACATA 1183
QY 321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
DB 1184 ATGGAAATTTGGAATTTAAATAATGTTTACTTCAACTTACCCTATCCAGAGTGAAGTTAAG 1243
QY 341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
DB 1244 ATCTGAAGGCCCTCAACCTGAAGGTGAAGCGGAGCGGTAGCCCTGGTTGGCAAC 1303
QY 361 AsnGlySerGlyLysSerThrValValGlnLeuGlnArgLeuTyrAspProAspAsp 380
DB 1304 AGTGGCTGTGGGAAAGACACAACTGTCCAGCTGCTGCAGAGGCTCTACGACCCCATAGAG 1363
QY 381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
DB 1364 GGCAGGTCTAGTATCAGCGGACAGGACATCAGGACCATCAATGTGAGGTATCTGCGGAA 1423
QY 401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnIle 420
DB 1424 ATCATTTGGGTGTGTAGTTCAGGAACCCGCTGTTTGCACCCAGATTGCCGAAACATT 1483
QY 421 LysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
DB 1484 CGTATGGCCGAGAAACCTGTCACCATGATGATAGAGAAAGCTGTCAAGGAGCCAAAT 1543
QY 441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyLysGly 460
DB 1544 GCCTATGACTTCATCATGAACTGCCCCCAAAATTTAAACACCCCTGGTTGGTGAGAGGG 1603
QY 461 AlaGluMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
DB 1604 GCGCAGTGTAGTGGGACAGAAACAGAGATCGCCATTCGCCGGCCCTGGTCCGCAAC 1663
QY 481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
DB 1664 CCCAAGATCCTTTTGTGGATGAGGCCACGTCAGCTTGGACACAGAAAGCGGCGTG 1723
QY 501 ValGlnAlaAlaLeuGluLys 507
DB 1724 GTTCAGCGCGCTCTGGATAG 1744

RESULT 2

US-09-450-105-1

; Sequence 1, Application US/09450105

; Patent No. 6169166

; GENERAL INFORMATION:

; APPLICANT: Kimberly Anne Brun

; APPLICANT: Richard James Chenery

; APPLICANT: Harma Ellens

; APPLICANT: John Anthony Feild

; APPLICANT: Lin Yue

; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES

; FILE REFERENCE: GP-50008-D1

; CURRENT APPLICATION NUMBER: US/09/450,105

; EARLIER APPLICATION NUMBER: 1999-11-29

; EARLIER FILING DATE: 1998-07-22

; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 4233
; TYPE: DNA
; ORGANISM: HOMO SAPIENS
US-09-450-105-1

Alignment Scores:

Pred. No.: 2,88e-179 Length: 4233
Score: 1576.00 Matches: 308
Percent Similarity: 75.90% Conservatives: 92
Best Local Similarity: 58.44% Mismatches: 105
Query Match: 60.80% Indels: 22
DB: 4 Gaps: 4

US-09-873-409-8 (1-514) x US-09-450-105-1 (1-4233)

```
QY 1 MetIleuGlyIleuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 170 ATGGCTCTGGGAACCTCGCTGCTATCATCCACGGAACCTGCTCCCTCTGATGCTG 229
QY 21 ValIleuGlyLeuMetSerAspAsnLeuLeu----- 30
Db 230 GTTTCGGATACATGACAGATAGTTTATCCCAAGCAGAGACCCGCAATCTGCGGAGCTT 289
QY 31 -----SerGlyCysLeuValGlnThrAsnThrTy+SerPheArgLeu----- 45
Db 290 ACTAATCAAGTGAATCAACAGTACACAGCCGTCAGCGACAGCAGTCTGGAGGAGGAC 349
QY 46 -----ThrLeuTyTrpValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db 350 ATGGCCATGTACGCCCTACTATTACAGGGCATTTGGTGGCGGTGCTCATCGTTGCTTAC 409
QY 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrIleArgIleGlyGln 81
Db 410 ATCCAGGTTTCACTTTGGTGGCGGAGTGGGAGACAAATACACAGATTAGGCGAGAAG 469
QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db 470 TTTTTCATGCCATCATGAATCAGGAGATAGGCTGGTTTACGTGATGACGCTGGGGAG 529
QY 102 LeuAsnThrArgMetThr---AspIleAspIleSerAspGlyIleGlyAspIle 120
Db 530 CTCAACACCCGGCTCACAGATGACGCTCCAAATAATGACCGAATGGTGACAACTT 589
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db 590 GGAATGTTCTTTTCAGTCCATAACGACATTTTTCAGCCGGTTTATATAGGATTTATAAGT 649
QY 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db 650 GGTGGAAGCTAACCTTGTAATTTTGGCGTCCAGCCCTCTTATTTGGTGTGCTATCGCC 709
QY 161 AlaCysSerArgMetValIleThrSerLysGluLeuSerAlaTySerLysAla 180
Db 710 ATCTGGCAAGAGTACTGACTCTATTACTAATAAGGAACCTCCAGGCTTATGCAAGCT 769
QY 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db 770 GGAGCAGATTGCCGAAGAAGCTTTAGCAGCCATCAGAACTGTGATGGTTGGAGGACAA 829
QY 201 GluLysGluLeuGlnArgTyThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
Db 830 AAGAAGGAACCTTGAAGGTACAATAAAAAATTTAGAAGAGCTAAAGAGTTGGCATAAG 889
QY 221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyPhePheMetAsnGlyTyTrp 240
Db 890 AAGACCATCAGGCCCAACATTTCCATAGGTATTGCCCTACCTGTTGGTCTATGCTCTAT 949
QY 241 GlyLeuAlaPheTrpTyTrpGlyThrSerLeuIleLeuAsnGlyGluProGlyTyThrIle 260
Db 950 GCACCTGGCATCTCTGGTATGGGACCTCTGCTCTCTCAAAATGAA-----TATTCTATT 1003
```

```
QY 261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyTrpCysIleGlyAlaAla 280
Db 1004 GGACAAGTCTTACCGCTCTTCTCTCTATTTATTGGGCACTTTTCAGTATTGGCATTTA 1063
QY 281 ValProHisPheGluThrPheAlaIleAlaArgGlyValAlaPheHisIlePheGlnVal 300
Db 1064 GCCCAAAACATAGAAGCCTTTGCAATGCAAGAGGGGAGCCCTTAATAATCTTCAAGATA 1123
QY 301 IleAspLysProSerIleAspAsnPheSerThrAlaGlyTyTrpLysProGluSerIle 320
Db 1124 ATGTATAATGAGCCCAAGCATTCACAGCTTCTCAACCAAGGACACAAACACAGACATTA 1183
QY 321 GluGlyThrValGluPheLysAsnValSerPheAsnTyTrpProSerArgProSerIleLys 340
Db 1184 ATGGGAATTTGGAATTTAAAAATGTTTACTTCAATACCCCATCACGAAAGTGAAGTTAAG 1243
QY 341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
Db 1244 ATCTTGAAGGGCTCAACCTGAAGGTGAAGAGCGGACAGCGTAGCCCTGGTTGGCAAC 1303
QY 361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyTrpAspProAspAsp 380
Db 1304 AGTGGCTGTGGGAAAGACAACTGTCCAGCTGCTGCAGAGGCTCTACGCCCCCATAGAG 1363
QY 381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyTrpArgAsp 400
Db 1364 GCGCAGGTCTAGTATCGACGACAGGACATCAGGACCATCAATGTGAGGTATCTGCGGAA 1423
QY 401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
Db 1424 ATCATTTGGGGTGGTGTAGTCAGGAACCCGTGCTTTGCCACCACGATTCGCCGAAACAT 1483
QY 421 LysTyTrpGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
Db 1484 CGCTATGGCGGAAACCTCACCATGATGATAGAGAAGCTCTCAGGAGAGCCAT 1543
QY 441 AlaTyTrpAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGly 460
Db 1544 GCCTATGACTTCATCATGAACTGCCCCCAAAATTTAACACCCCTGTTGTGTGAGAGAGG 1603
QY 461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db 1604 GCGCAGCTAGTGGGGACAGAAACAGAGAGTCCCATTTGCCCGGCCCTGGTCCGCAAC 1663
QY 481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db 1664 CCAAGATCTTTTGTGATGAGGCCCATCTGACCTTTGGACACAGAAAGCGGCGGTG 1723
QY 501 ValGlnAlaAlaLeuGlyLys 507
Db 1724 GTTCAGCGGCTCTGGATAAG 1744
```

RESULT 3

US-08-752-447-1

; Sequence 1, Application US/08752447

; Patent No. 5994088

; GENERAL INFORMATION:

; APPLICANT: Mechtner, Eugene

; APPLICANT: Roninson, Igor B

; TITLE OF INVENTION: Methods and Reagents for Preparing and

; TITLE OF INVENTION: Using Immunological Agents Specific for P-glycoprotein

; NUMBER OF SEQUENCES: 2

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff Ltd.

; STREET: 300 South Wacker Drive, Seventh Floor

; CITY: Chicago

; STATE: Illinois

; COUNTRY: USA

; ZIP: 60606

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/752,447

FILING DATE: 15-NOV-1996

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: NO. 5994088nan, Kevin E

REGISTRATION NUMBER: 35,303

REFERENCE/DOCKET NUMBER: 95,1121

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312-913-0001

TELEFAX: 312-913-9808

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 4669 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: 5'UTR

LOCATION: 1..424

FEATURE:

NAME/KEY: CDS

LOCATION: 425..4264

FEATURE:

NAME/KEY: 3'UTR

LOCATION: 4265..4669

US-08-752-447-1

Alignment Scores:

Pred. No.: 3 4e-179 Length: 4669
Score: 1576.00 Matches: 305
Percent Similarity: 77.04% Conservative: 101
Best Local Similarity: 57.87% Mismatches: 99
Query Match: 60.80% Indels: 22
DB: 2 Gaps: 5

US-09-873-409-8 (1-514) x US-08-752-447-1 (1-4669)

```
Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 575 ATGGTGGGGAACCTTTGGCTGCCATCATCCATGGGCTGGACTTCCTCTCATGATGCTG 634
Qy 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
Db 635 GTGTTGGAGAAATGCACAGATATCTTTGCAAAATGCAGAAATTTAGAAATCTCATGTCA 694
Qy 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
Db 695 AACATCACTAATAGAGTGATATCAATGATACAGGTTCTTCATGAATCTGGAGGAAGAC 754
Qy 44 -----ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db 755 ATGACAGGTATGCTTATTATACAGTGAATTTGGTGTGGGCTGGTGTGCTTAC 814
Qy 62 IleGlnIleSerLeuTyrPheIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db 815 ATTCAGGTTTCATTGTTGGTGTGGCAGCTGGAAGACAAATACACAAATTTAGAAACAG 874
Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101
Db 875 TTTTTCATGCTTATATCCGACAGAGATAGCTGGTTTGTATGTCACGATGTTGGGAG 934
Qy 102 LeuAsnThrArgMetThr-----AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db 935 CTTAACACCCGACITACAGATGATGCTCTCCAGATTAATGAAGAAATTTGGTGACAAAT 994
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db 995 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCATCTGGGTTTATAGTAGGATTTACACGT 1054
```

```
Qy 141 GlyTyrPlyLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db 1055 GGTGTGAAGCTAAACCTTTGTGATTTTGGCCATCAGCTCTGTTCTTGACGTGTCAGTCT 1114
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db 1115 GTCGTGGCAAGATACCTATCTTCTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 1174
Qy 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db 1175 GGAGCAGTAGCTGAAGAGCTCTTGGCAGCAATTTAGAACTGTGATTTGCATTTGGAGCAA 1234
Qy 201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
Db 1235 AAGAAAGAACTTGAAGGTACAAACAAATTTAGAAAGCTTAAAGAAATTTGGATAAAG 1294
Qy 221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
Db 1295 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTCTCTGCTGATCTATGATCTTAT 1354
Qy 241 GlyLeuAlaPheTyrTyrGlyThrSerLeuIleAsnGlyGluProGlyTyrThrIle 260
Db 1355 GCTGTGGCTTCTGTTATGGGACCACTTGTCTCTCAGGGAA-----TATCTATT 1408
Qy 261 GlyThrValLeuAlaValPheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
Db 1409 GGACAGTACTCACTGTATTTCTTTCTGTTAATTTGGGGCTTTTAGTGTGGCAGGCA 1468
Qy 281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
Db 1469 TCTCCAAGCATTTGAAGCATTTGCAAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1528
Qy 301 IleAspLysIleProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
Db 1529 ATTGATAAATGAAGCAAGTATTGACAGCTATTGAAAGTGGGACCAACAGATAATATT 1588
Qy 321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
Db 1589 AAGGAAATTTGGAATTCAGAAATGTTCACTTCACTTACCAATCTCGAAAGAAAGTTAAG 1648
Qy 341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
Db 1649 ATCTGAAGGCTGCTGAACCTGAAGTGCAGAGTGGCAGACGGTGGCCCTGGTGGAAAC 1708
Qy 361 AsnGlySerGlyLysSerThrValGlnLeuGlnArgLeuTyrAspProAspAsp 380
Db 1709 AGTGCTGTGGGAAGACACACAGCTCAGCTGATGACAGAGGCTCTATGACCCACAGAG 1768
Qy 381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
Db 1769 GGGATGTCAGTGTGATGGACAGGATATTAGGACCAATAAATGTAAGGTTTCTACGGAA 1828
Qy 401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnIle 420
Db 1829 ATCATTTGGTGTGGTGCAGGAACCTGTATTGTTGCCACCACGATAGCTGAAACATT 1888
Qy 421 LysTyrGlyArgAspValThrAspGluGluMetClnuArgAlaAlaArgGluAlaAsn 440
Db 1889 CGCTATGGCGTGAATGTCCATGATGAGATGAGAAAGCTGTGAGAAAGCTGTGAGAGAGG 1948
Qy 441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyLysGly 460
Db 1949 GCCTATGACTTTTATCATGAACCTGCCTCATATAATTTGACACCTGTTGGTGGAGAGAGG 2008
Qy 461 AlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db 2009 GCCCAGTTGATGTGGGCAAGACAGAGGATGCCATTCAGCTGCGCTGGTTTCGCAAC 2068
Qy 481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db 2069 CCCAAGATCTCTCTGCTGATGAGGCCACCTGACGCTTGGACACAGAAAGGACGAGT 2128
Qy 501 ValGlnAlaAlaLeuGluLys 507
```



```
QY 381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
Db 1769 GGGATGGTCAGTGTGGACAGGATATTAGGACCAATAAATGTAAGGTTTCTACGGGAA 1828
QY 401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
Db 1829 ATCATTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1888
QY 421 LysTyrGlyArgAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsn 440
Db 1889 GCGTATGGCCGTGAAATGTCACCATGGATGAGATTGAGAAAGCTGCAAGGAGGCCAAT 1948
QY 441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
Db 1949 GCCTATGACTTATCATGAACCTGCTCATATAATTTGACACCCCTGGTGGAGAGAGGG 2008
QY 461 AlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaAlaLeuValArgAsn 480
Db 2009 GCCCAGTTGAGTGGTGGCAGAGCAGAGGATGCCATTGACGTGCCCTGGTTGCGCAAC 2068
QY 481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db 2069 CCCAAGATCTCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 2128
QY 501 ValGlnAlaAlaLeuGluLys 507
Db 2129 GTTCAGGTGGCTCTGGATAG 2149
RESULT 5
US-08-181-471-2
; Sequence 2, Application US/08181471
; Patent No. 5641508
; GENERAL INFORMATION:
; APPLICANT: Li, Lingna
; APPLICANT: Lishko, Valeryi K.
; TITLE OF INVENTION: METHOD FOR DELIVERING BENEFICIAL
; TITLE OF INVENTION: COMPOSITIONS TO HAIR FOLLICLES
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Thomas Fitting
; STREET: 12526 High Bluff Drive, Suite 300
; CITY: San Diego
; STATE: CA
; COUNTRY: USA
; ZIP: 92130
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/181,471
; FILING DATE: 13-JAN-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/041,553
; FILING DATE: 02-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: ANT0029P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-792-3680
; TELEFAX: 619-792-8477
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4646 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
```

```
; ANTI-SENSE: NO
; FEATURE: CDS
; NAME/KEY: CDS
; LOCATION: 425..4267
US-08-181-471-2
Alignment Scores:
Pred. No.: 1,79e-178 Length: 4646
Score: 1570.00 Matches: 305
Percent Similarity: 76.85% Conservative: 100
Best Local Similarity: 57.87% Mismatches: 100
Query Match: 60.57% Indels: 22
DB: 1 Gaps: 5
US-09-873-409-8 (1-514) x US-08-181-471-2 (1-4646)
QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 575 ATGGTGGTGGGAACCTTGGCTGCCATCATCATCATCATCATCATCATCATCATCATCATCAT 634
QY 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
Db 635 GTGTTTGGGAAATGACAGATATCTTTGCAAAATGCAGGAAATTTAGAAAGATCTGATGTCA 694
QY 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
Db 695 AACATCACTAATAGAAAGTATATCAATGATATACAGGCTTCTTCATGAATCTGGAGGAAGAC 754
QY 44 -----ArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db 755 ATGACCAGGTATGCCATTATTATACAGTGGAAATTTGGTGGGTGCTGGTGTGCTGTAC 814
QY 62 IleGlnIleSerLeuTrpIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db 815 ATTCAGTTTCATTTTGGTGGCTGGCAGCTGGAGACAAATACACAAATTAGAAACAG 874
QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101
Db 875 TTTTTCATGCTATAATATGCAGAGGATAGGCTGGTTGATGTGCACGATGTGGGAG 934
QY 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db 935 CTTAACAACCCGACTTACAGATGATGCTCTTAAGATTAAATGAAGTTATTGGTGCAAAAT 994
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db 995 GGAATGTTCTTCAGTCAATGGCAACATTTTCACCTGGGTTATAGTAGGATTACAGCT 1054
QY 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db 1055 GGTTGGAAGCTAACCTTGTGATTTTGGCCATCAGTCTCTTCTGGACTGTGAGCTGCT 1114
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db 1115 GTCTGGGCAAGATACACTTCTTACTGTATAAAGAACTCTTAGCGTATGCAAAAGCT 1174
QY 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db 1175 GGAGCAGTAGCTGAGAGAGTCTTGGCAGCAATTAGAATCTGTGATTCATTTGAGAGCAA 1234
QY 201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
Db 1235 AAGAAAGAACTTCAAGAGGTACAAACAAATTTAGAAAGAGCTAAAGAAATTTGGATAAAG 1294
QY 221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
Db 1295 AAAGCTATTACAGCCCAATATTTCTATAGTGGTGGCTTTCTCTGCTGATCTATCATCTAT 1354
QY 241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
Db 1355 GCTCTGGCCCTTCGGTATGGACACCTGGTGGTCTCTCAGGGGAA-----TATTCTATT 1408
QY 261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
```



```
Db 1409 GGACAGTACTCACTGATCTCTTCTGTTAATTTGGGGCTTTAGTGTGGACAGGCA 1468
Qy 281 ValProHisPheGluThrPheAlaIleAlaArgGlyValAlaPheHisIlePheGlnVal 300
Db 1469 TCTCCAGCATTTGAACATTTGCAATGCAAGAGGAGCAGCTATGAAATCTTCAAGATA 1528
Qy 301 IleAppLysPheProSerIleAspPheSerThrAlaGlyTyrLysProGluSerIle 320
Db 1529 ATTGATAAAGCCAGTATTGACAGCTATTGACAGCTATTGACAGCTGGGCACAAACAGATATAT 1588
Qy 321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
Db 1589 AAGGGAATTTGGAATTCAGAAATGTTCACTTCAGTATCCCATCTCGAAAGAGTTAAG 1648
Qy 341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
Db 1649 ATCTTGAAGGGCTGAACTGAAAGTGCAGAGTGGCAGAGCGTGGCCCTGGTGGAAAC 1708
Qy 361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
Db 1709 AGTGGCTGTGGGAAGAGCACACAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAG 1768
Qy 381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
Db 1769 GGGATGCTCAGTGTGATGGACAGAGTATTAGGACCATAAATGTAAGGTTTCTACGGGAA 1828
Qy 401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
Db 1829 ATCAITGGTGTGGTGTGAGTCAAGAACTGTATTGTTGCCACACAGATAGCTGAAACATT 1888
Qy 421 LysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
Db 1889 CGCTATGGCGGTGAAATGTCACCATGGATGAGATGAGAAAGCTGTCAAGGAAGCCAAAT 1948
Qy 441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
Db 1949 GCCTATGACTTTATCATGAACCTGCCTCATATAATTTGACACCCCTGGTGGAGAGAGGG 2008
Qy 461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db 2009 GCCAGTTGAGTGTGGGAGAGAGAGAGATGCCATTGACGTCGCCCTGGTTCGCAAC 2068
Qy 481 ProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db 2069 CCAAGATCTCTCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAGCAGTG 2128
Qy 501 ValGlnAlaLeuGluLys 507
Db 2129 GTTCAGGTGGCTCTGGATAAG 2149
```

RESULT 6

```
US-08-793-610-5
; Sequence 5, Application US/08793610
; Patent No. 5858744
; GENERAL INFORMATION:
; APPLICANT: BAUM, Christopher
; APPLICANT: STOCKING-HARBERS, Carol
; APPLICANT: OSTERTAG, Wolfram
; TITLE OF INVENTION: RETROVIRAL VECTOR HYBRIDS AND THE USE THEREOF
; TITLE OF INVENTION: FOR GENE TRANSFER
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nikaido, Marmelstein, Murray & Oram LLP
; STREET: 655 Fifteenth Street N.W. Suite 330
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/793,610
; FILING DATE: 07-MAR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 44 31 973.8
; FILING DATE: 08-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 195 03 952.1
; FILING DATE: 07-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP95/03175
; FILING DATE: 10-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Berman, Richard J.
; REGISTRATION NUMBER: 39,105
; REFERENCE/DOCKET NUMBER: P1614-7007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)638-5000
; TELEFAX: (202)638-4810
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6505 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA
US-08-793-610-5

Alignment Scores:
Pred. No.: 3,17e-178 Length: 6505
Score: 1570.00 Matches: 305
Percent Similarity: 76.85% Conservative: 100
Best Local Similarity: 57.87% Mismatches: 100
Query Match: 60.57% Indels: 22
DB: 2 Gaps: 5

US-09-873-409-8 (1-514) x US-08-793-610-5 (1-6505)
Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyValAlaCysLeuProLeuMetProLeu 20
Db 1967 ATGGTGGTGGGAACCTTGGCTGCCATCATCCATGGGGCTGGACTTCTCTCATGATGCTG 2026
Qy 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
Db 2027 GTGTTTGGAGAAATGACAGATATCTTTGCAAAATGCAGGAAATTTAGAGATCTGATGTCA 2086
Qy 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
Db 2087 AACATCACTAATAGAGTATGATATCAATGATACAGGGGTTCTTCATGAATCTGGAGGAAGAC 2146
Qy 44 -----ArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db 2147 ATGACCAAGTATGCCATTATTATACAGTGAATTTGGTGGGTGCTGGTGTGCTGTAC 2206
Qy 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db 2207 ATTCAGTTTTCATTTTGGTGGCTGGCAGCTGGAGACAAATACACAAATTAGAAAACAG 2266
Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101
Db 2267 TTTTTCATGCTATATATGCGACAGGAGATAGGCTGTTTTCATGTGCACCATGTTGGGGAG 2326
Qy 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db 2327 CTTAACACCCGACTACATGATGATGCTCTTAAGATTAAAGATTATTTGGTGCAAAAT 2386
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db 2387 GGAATGTTCTTCAGTCAATGGCAACATTTTTCACATGGGGTTTATAGTAGGATTTACACGT 2446
Qy 141 GlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
```


Db 2447 GGTTGGAAGCTAACCCCTGTGTCATTTTGGCCATCAGTCTCTGTCTTGACACTGTGACGTGCT 2506
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerIysGluLeuSerAlaTyrSerIysAla 180
Db 2507 GTCTGGCAAGATACATCTTCTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 2566
Qy 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db 2567 GGAGCACTAGCTGAAGAGGCTTGGCAGCAATAGAACTGTGATGCTTTGGAGGACAA 2626
Qy 201 GluIysGluLeuGlnArgTyrThrGlnAsnLeuIysAspPheGlyIleIys 220
Db 2627 AAGAAAGAACTTGAAGGTACACAAAATTTAGAGAACTAAAGATTTGGATAAAG 2686
Qy 221 ArgThrIleAlaSerIysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
Db 2687 AAGCTATTACAGCCAAATTTCTATAGGTGCTCTTCTCTGCTGATCTATGCAATCTAT 2746
Qy 241 GlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
Db 2747 GCTCTGGCCTTCTGTATGGACCACTTGGTCTCTCAGGGAA-----TATCTATT 2800
Qy 261 GlyThrValLeuAlaValPheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
Db 2801 GGCAAGTACTCTACTGTATCTTCTCTGATTAATTGGGGCTTTTAGTGTGGACAGCA 2860
Qy 281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
Db 2861 TCTCCAAGCAATTGAAGCATTTGCAAAATGCAAGAGGACGCTTATGAATCTTCAAGATA 2920
Qy 301 IleAspIysIysProSerIleAspAsnPheSerThrAlaGlyTyrIysProGluSerIle 320
Db 2921 ATTGATTAATAGCCAGTATTGACAGTATTCGAGAGTGGGCACAAACAGATAATT 2980
Qy 321 GluGlyThrValGluPheIysAsnValSerPheAsnTyrProSerArgProSerIleIys 340
Db 2981 AAGGAAATTTGGAATTCAGAAATGTTCACTTCACTTACCATCTCGAAAGAAAGTTAAG 3040
Qy 341 IleLeuIysGlyLeuAsnLeuArgIleIysSerGlyGluThrValAlaLeuValGlyLeu 360
Db 3041 ATCTTGAAGGCGCTGAACCTGAAGGTGCAGAGTGGCAGACGGTGGCCCTGTGTTGAAAC 3100
Qy 361 AsnGlySerGlyIysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
Db 3101 AGTGGCTGTGGAAAGACACACACAGTCCAGCTGATGACAGGCTCTATGCCCCACAGAG 3160
Qy 381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
Db 3161 GGGATGTCAGTGTGTGATGGACAGGATATTAGGACCAATAATGTAAGGTTTCTACGGAA 3220
Qy 401 HistGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
Db 3221 ATCAATGGTGTGGTGTGAGTACAGAACCTGTATGTTGCCACCGATAGCTGAAACATT 3280
Qy 421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
Db 3281 CGCTATGGCCGTGAATATGTCACCATCGATGATGAGATGAGAAGCTGTCAAGAAAGCCAT 3340
Qy 441 AlaTyrAspPheIleMetGluPheProAsnIysPheAsnThrLeuValGlyGluIysGly 460
Db 3341 GCCTATGACTTTTATCATGAAACTGCCTCATAAATTTGACACCCCTGGTTGGAGAGAGGG 3400
Qy 461 AlaGlnMetSerGlyGlyGlnIysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db 3401 GCCCAGTTGAGTGTGGCAGAAAGCAGAGGATCGCCATTGACAGTGGCCCTGTGTTGCAAC 3460
Qy 481 ProIysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerIysSerAla 500
Db 3461 CCCAAGATCCCTCTCTGCTGGATGAGGCCACGTCAGCCCTTGGACAGAAAGCCAGCATG 3520
Qy 501 ValGlnAlaAlaLeuGluIys 507
Db 3521 GTTCAGGTGGCTCTGGATAAG 3541

RESULT 7

US-08-793-610-6
; Sequence 6, Application US/08793610
; Patent No. 5858744
; GENERAL INFORMATION:
; APPLICANT: BAUM, Christopher
; APPLICANT: STOCKING-HARBERS, Carol
; APPLICANT: OSTERTAG, Wolfram
; TITLE OF INVENTION: RETROVIRAL VECTOR HYBRIDS AND THE USE THEREOF
; TITLE OF INVENTION: FOR GENE TRANSFER
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Nikaido, Marmelstein, Murray & Oram LLP
; STREET: 655 Fifteenth Street N.W. Suite 330
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/793,610
; FILING DATE: 07-MAR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 44 31 973.8
; FILING DATE: 08-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 195 03 952.1
; FILING DATE: 07-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP95/03175
; FILING DATE: 10-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Berman, Richard J.
; REGISTRATION NUMBER: 39,105
; REFERENCE/DOCKET NUMBER: P1614-7007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)638-5000
; TELEFAX: (202)638-4810
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9318 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA
US-08-793-610-6
Alignment Scores:
Pred. No.: 5,85e-178 Length: 9318
Score: 1570.00 Matches: 305
Percent Similarity: 76.85% Conservative: 100
Best Local Similarity: 57.87% Mismatches: 100
Query Match: 60.57% Indels: 22
DB: 2 Gaps: 5
US-08-793-409-8 (1-514) x US-08-793-610-6 (1-9318)
Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 1926 ATGGTGGTGGAACTTGGCTGCCATCATCCATGGGCTGGACTTCTCTCATGATGCTG 1985
Qy 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
Db 1986 GTGTTGGAGAAATGACAGATATCTTTGCAAAATGCAGGAATTTAGAGATCTGATGTC 2045
Qy 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
Db 2046 AACATCACTAATAGAGTGATATCAATGATACAGGGTTCCTTCATGAATCTGGAGGAAGAC 2105

QY 44 -----ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuLeuPheGlyTyr 61
Db 2106 ATGACCAGGTATGCCCTATTATTACAGTGAATTTGGTGGGTGCTGGTGTGCTTAC 2165
QY 62 IleGlnIleSerLeuTyrPheLeuThrAlaAlaArgGlnThrIleArgIleGlyGln 81
Db 2166 ATTCAGGTTTCAATTTGGTGGCTGGAGGAGCAATATACCAAAATTAAGAAACAG 2225
QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101
Db 2226 TTTTTCATGCTATATGCGACAGGAGATAGCTGTGTGATGTGCACGATGTGGGAG 2285
QY 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db 2286 CTTAACACCCGACTTACAGATGATGCTCTCAAGATTAAATGAAGTTATTGGTGACAAAAT 2345
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db 2346 GGAATGTTCTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGGATTTACACGT 2405
QY 141 GlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db 2406 GGTGGAAGCTAACCTTGTGATTTGGCCATCAGTCTCTGTTCTTGGACTGTGAGTGTCT 2465
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db 2466 GTCTGGGCAAGATACACTATCTTCTATGATAAAGAACTCTTAGCGTATGCAAAAGCT 2525
QY 181 GlyAlaValAlaGluLeuValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db 2526 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATATAGAACTGTGATTCATTTGGAGACAA 2585
QY 201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
Db 2586 AGAAAGAACTTGAAGAGGTACAAACAAATTTAGAAAGCTAAAGAAATTTGGGATAAG 2645
QY 221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
Db 2646 AAAGCTATTACGCCAATTTCTATAGTGTGCTTCTCTGCTGATCATGATCTTAT 2705
QY 241 GlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
Db 2706 GCTCTGGCTTCTGGTATGGACCACTTGGTCTCTCAGGGAA-----TATTTCTATT 2759
QY 261 GlyThrValLeuAlaValPhePheSerValIleHisSerTyrCysIleGlyAlaAla 280
Db 2760 GGACAAGTACTACTGATTTCTTTCTGTATTAAATTTGGGGCTTTTAGTGTGGACAGGCA 2819
QY 281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
Db 2820 TCTCCAAGCATTTGAAGCAATTTGCAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 2879
QY 301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
Db 2880 ATTGATAATAAGCAAGTATTGACAGTATTCGAAGTGGGACCAACACAGATAATATT 2939
QY 321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
Db 2940 AAGGGAAATTTGGAATTTGAAATGTTTCACTTCAGTTACCCTTCTGAAAGAGTTAAG 2999
QY 341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
Db 3000 ATCTGAGGGCTTGAACCTGAGAGTGCAGAGTGGGACAGCGTGGCCCTGGTTGGAAC 3059
QY 361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
Db 3060 AGTGGCTGTGGGAAGAGCACAAACAGTCCAGCTGATGCAGAGGCTCTATGACCCACAGAG 3119
QY 381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
Db 3120 GGGATGCTCAGTGTGTGATGGACAGGATATTAGGACCAATAATGTAAGGTTTCTACCGGAA 3179

QY 401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
Db 3180 ATCAITGGTGTGTGAGTCAGGAACCTGTATTGTTTGCACCACGATAGTGAACAAT 3239
QY 421 LysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
Db 3240 CGCTATGGCGGTGAATATGTCCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCAAT 3299
QY 441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGly 460
Db 3300 GCCTATGACTTTATCATGAACCTGCCTCATATAATTTGACACCCCTGTTGGAGAGAGGG 3359
QY 461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db 3360 GCCCAGTTTCAAGTGGGAGAGAGAGGATGCGCATTCACAGTCCCTGTTGTCGCAAC 3419
QY 481 ProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db 3420 CCAAGATCTCTGCTGGATGAGGCCACGTGACCTTGGACACAGAAAGCGAGCATG 3479
QY 501 ValGlnAlaAlaLeuGluLys 507
Db 3480 GTTCAGGTGGCTCTGGATAAG 3500
RESULT 8
US-08-784-649A-1
; Sequence 1, Application US/08784649A
; Patent No. 5830697
; GENERAL INFORMATION:
; APPLICANT: Sikic, Branimir I
; APPLICANT: Chen, Gang
; TITLE OF INVENTION: P-GLYCOPROTEIN MUTANT RESISTANT TO
; TITLE OF INVENTION: CYCLOSPORIN MODULATION
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Fish & Richardson
; CITY: Menlo Park
; STATE: CA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/784,649A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: Reg. No. 5830697 36,677
; REFERENCE/DOCKET NUMBER: 06037/007001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-322-5070
; TELEFAX: 415-854-0875
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4264 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-784-649A-1
Alignment Scores:
Pred. No.: 2,34e-178 Length: 4264
Score: 1568.50 Matches: 306
Percent Similarity: 77.04% Conservatives: 100
Best Local Similarity: 58.08% Mismatches: 98
Query Match: 60.51% Indels: 23
DB: 2 6

; LENGTH: 4264 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-784-649A-5

Alignment Scores:

Pred. No.: 2,346-178 Length: 4264
Score: 1568.50 Matches: 306
Percent Similarity: 77.04% Conservative: 100
Best Local Similarity: 58.06% Mismatches: 98
Query Match: 60.51% Indels: 23
DB: 2 Gaps: 6

US-09-873-409-8 (1-514) x US-08-784-649A-5 (1-4264)

```
Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 292 ATGGTGGTGGGAACCTTTGGCTGCATCATCCATGGGGCTGGACCTCTCTCATGATGCTG 351
Qy 21 ValLeuGlyLeuMetSerAsp-----AsnLeuIleSer 31
Db 352 GGTGTTGGAGAAATGACATATCTTTGCAATGCAAGGAAATTTAGAAATCTGATGCTCA 411
Qy 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
Db 412 AACATCACTAATAGAGATGATATCAATGATACAGGGTCTTCTCATGAATCTGGAGGAAGAC 471
Qy 44 -----ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db 472 ATGACCAGGTATGCCCTATTATTACAGTGAATTTGGTGGGGTGGCTGCTGTCTTAC 531
Qy 62 IleGlnIleSerLeuTrpIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db 532 ATTCAGGTTTCATTTGGTGGCTGGCAGCTGGAGAACAAATACACAAATTAGAAAACAG 591
Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101
Db 592 TTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTGATGTCACAGATGTTGGGAG 651
Qy 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db 652 CTTAACACCCGACTTACAGATGATGCTCCCAAGTAAATGAAGAAATGCTGACAAAATT 711
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db 712 GGA---ATGTTCCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGGATTTACACGT 768
Qy 141 GlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db 769 GGTGGAAGCTAACCTTGTGATTTGGCCATCAGTCTCTTGGACCTGTCAGCTGCT 828
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysLeuLeuSerAlaTyrSerLysAla 180
Db 829 GTCTGGGCAAGATACTATCTTCAATTAATGATAAAGAACTCTTAGCGTATGCAAAAGCT 888
Qy 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db 889 GGAGCAGTAGCTGAAGAGGCTTTGGCAGCAATTTAGAACTGTGATTCATTTGGAGACAA 948
Qy 201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspPheGlyIleLys 220
Db 949 AAGAAGAACTTGAAGGTACACAAAATTTAGAGAACTTAAGAATTTGGATAAG 1008
Qy 221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
Db 1009 AAAGCTATTACGCCAATATTTCTATAGGTGCTGCTTCTGCTGATCTATGATCTTAT 1068
Qy 241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
Db 1069 GCTCTGGCTTCTGGTATGGGACCACTTGGTCTCTCAGGGGAA-----TATTCTATT 1122
```

```
Qy 261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
Db 1123 GGACAAGTACTCACTGATATCTTTCTGTATTAAATGGGGCTTTTAGTGTGGACAGGCA 1182
Qy 281 ValProHisPheGlyThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
Db 1183 TCTCCAGCATTTGAAGCATTTGCCAAATGCAAGAGGAGCAGCTTATGAATCTTCAAGATA 1242
Qy 301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
Db 1243 ATTGATAATAAGCCAAAGTATTGACAGCTATTGCAAGAGTGGGCAACACAGATAATATT 1302
Qy 321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
Db 1303 AAGGGAAATTTGGAATTTCAAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAGTTAAG 1362
Qy 341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
Db 1363 ATCTTGAAGGGCTGAACTGAAGTGCAGATGGGCACAGCGTGGCCCTGTTGGAAAC 1422
Qy 361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
Db 1423 AGTGGCTGTGGGAAGAGCACAAACAGTCCAGCTGATGCAGAGGCTCTATGACCCACAGAG 1482
Qy 381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
Db 1483 GGGAGTGTGAGTGTGATGCAGCAGGATATTAGGACCATTAATGTAAGGTTTCTACGGGAA 1542
Qy 401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
Db 1543 ATCAATTTGGTGTGAGTGCAGAACTGTATTTGTTGGCCACACGATAGCTGAAACATTT 1602
Qy 421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
Db 1603 CGCTATGCGCGTGAAATGTCACCATGGATGAGATTGAGAAAGCTCTCAAGGAAGCCAAT 1662
Qy 441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
Db 1663 GCCTATGATCTTATCATGAAACCTGCCTCATAAATTTGACACCCCTGTTGGAGAGAGGG 1722
Qy 461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db 1723 GCCAGTTGAGTGGTGGGAGAGCAGAGATCGCCATTGACGTGCCCTGTTGTTGCGAAC 1782
Qy 481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db 1783 CCCAAGATCTCTGCTGATGAGGCCACGTCAGCTTGGACACAGAAAGCAGAGCATG 1842
Qy 501 ValGlnAlaAlaLeuGluLys 507
Db 1843 GTTCAGGTGGCTCTGGATAAG 1863
RESULT 10
5206352-3
; Patent No. 5206352
; APPLICANT: Roninson, Igor B.; Pastan Ira H.; Gottesman,
; Michael M.
; TITLE OF INVENTION: COMPOSITIONS FOR CLONES CONTAINING DNA
; SEQUENCES ASSOCIATED WITH MULTIDRUG RESISTANCE IN HUMAN CELLS
; NUMBER OF SEQUENCES: 4
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/622,836
; FILING DATE: 24-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 892,575
; FILING DATE: 01-AUG-1986
; APPLICATION NUMBER: 845,610
; FILING DATE: 28-MAR-1986
; SEQ ID NO:3:
; LENGTH: 4669
5206352-3
Alignment Scores:
```



```
Pred. No.: 3,14e-178 Length: 4669
Score: 1568.00 Matches: 305
Percent Similarity: 76.66% Conservative: 99
Best Local Similarity: 57.87% Mismatches: 101
Query Match: 60.49% Indels: 22
DB: Gaps: 5

US-09-873-409-8 (1-514) x 5206352-3 (1-4669)

QY 1 MetIleuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
DB 575 ATGGTGGTGGAACTTTGGTGGCCATCATCGGGCTGGACTTCTCTCATGATGCTG 634
QY 21 ValLeuGlyLeuMetSerAsp-----AsnLeuIleSer 31
DB 635 GTGTTGGAGAAATGACAGATATCTTTGCAATGCAGGAATTTAGNAGATCTGATGTC 694
QY 32 GlyCysLeuValGlnThr-----AsnThrTySerPhePhe----- 43
DB 695 AACATCACTAATAGAGTATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 754
QY 44 -----ArgLeuThrLeuTyThrValGlyIleGlyValAlaAlaLeuIlePheGlyTy 61
DB 755 ATGACAGGTATGCTTATATACAGTGAATTTGGTGGTGGTGGTGGTGGTGGTGGT 814
QY 62 IleGlnIleSerLeuTrpIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
DB 815 ATTCAGGTTTCATTGGTGGCTGGAGCTGGAGACAAATACACAAAATTTAGAAAACAG 874
QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
DB 875 TTTTTCATGCTAATATCGACAGGAGATAGGCTGTTGATGTCACGATGTTGGGAG 934
QY 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
DB 935 CTTAAACACCGACTTACAGATGATGCTCTTAAGATTAATCAAGTTATTTGGTGAACAAAT 994
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
DB 995 GGAATGTTCTTCAGTCAATGGCAACATTTTCACTGGGTTTATAGTAGATTTACAGCT 1054
QY 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
DB 1055 GGTGGAAAGCTAACCTTGTGATTTGGCCATCAGTCTGCTTCTTGGAGCTGCTGCTGCT 1114
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTySerLysAla 180
DB 1115 GTCTGGGCAAGATACTATCTTCAATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 1174
QY 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
DB 1175 GGAGCAGTACTGAGAGGCTTTGGAGCAATAGAACTGTGATGCTTTGGAGGACAA 1234
QY 201 GluLysGluLeuGlnArgTyThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
DB 1235 AAGAAAGAACTTGAAGGTACAAACAAATTTAGAAAGCTAAAGAAATTTGGGATAAG 1294
QY 221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyThrPheMetAsnGlyThrTy 240
DB 1295 AAAGCTATTACAGCAATATTTCTATAGGTGCTGCTTCTGCTGATGCTATGCTATCTTAT 1354
QY 241 GlyLeuAlaPheTrpTyThrSerLeuIleLeuAsnGlyGluProGlyTyThrIle 260
DB 1355 GCTCTGGCTTCTGATGAGGACCACTTGGTCTCTCAGGGNA-----TATCTATT 1408
QY 261 GlyThrValLeuAlaValPheSerValIleHisSerTySerTyCysIleGlyAlaAla 280
DB 1409 GGACAACTACTCTGATTTCTTCTGATTAATTTGGGCTTTTAGTGTGGACAGCA 1468
QY 281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaPheHisIlePheGlnVal 300
DB 1469 TCTCCAGCATTTGAAGCATTTGCAATGCAAGAGGAGCTTATGAATCTTCAAGATA 1528
```

```
QY 301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyThrLysProGluSerIle 320
DB 1529 ATTGATAATAAGCAAGTATTGACAGCTATTGCAAGAGTGGGCAAAACCCAGATAATATT 1588
QY 321 GluGlyThrValGluPheLysAsnValSerPheAsnTyProSerArgProSerIleLys 340
DB 1589 AAGGGAAATTTGGAAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAGAAAGTTAAG 1648
QY 341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
DB 1649 ATCTTGAGGGCTGAACCTGAAGTGCAGTGGCAGCAGCGTGGCCCTGGTTGGAAC 1708
QY 361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyThrAspProAspAsp 380
DB 1709 AGTGGCTGTGGGAAGAGCACACAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAG 1768
QY 381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyArgAsp 400
DB 1769 GGGATGCTGAGTGTGATGACAGAGATATTAGGACCAATAATGTAAGGTTTCTACCGGAA 1828
QY 401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
DB 1829 ATCATTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1888
QY 421 LysTyArgLysArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
DB 1889 CGCTATGGCGTGAATAATGTCACCATGATGATGATGATGATGATGATGATGATGATGAT 1948
QY 441 AlaTyArgPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGly 460
DB 1949 GCCTATGACTTTATCATGAAACTGCTCAATAATTTGACACCTGTTGGAGAGAGAGGG 2008
QY 461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
DB 2009 GCCAGTTGAGTGGTGGGAGAGAGAGAGATGCCATGACGCTGCTGCTGCTGCTGCTGCT 2068
QY 481 ProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
DB 2069 CCCAAGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2128
QY 501 ValGlnAlaAlaLeuGluLys 507
DB 2129 GTTCAGGTGCTCTGTAAG 2149

RESULT 11
US-08-583-276-18
; Sequence 18, Application US/08583276
; Patent No. 5837536
; GENERAL INFORMATION:
; APPLICANT: McDonagh, Kevin T.
; APPLICANT: Nienhuis, Arthur
; APPLICANT: Tolstoshev, Paul
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HUMAN
; TITLE OF INVENTION: MULTIDRUG RESISTANCE GENES AND IMPROVED
; TITLE OF INVENTION: SELECTION OF CELLS TRANSDUCED WITH SUCH GENES
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan,
; ADDRESSEE: Cecchi & Stewart
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: DW4.V2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/583,276
; FILING DATE: 05-JAN-1996
; CLASSIFICATION: 435
```


STREET: Lilly Corporate Center
CITY: Indianapolis
STATE: Indiana
COUNTRY: USA
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/612,521
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Hamilton, Amy E
REGISTRATION NUMBER: 33,894
REFERENCE/DOCKET NUMBER: X-9693
TELEPHONE: 317-276-3169
TELEFAX: 317-276-1294
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4224 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 1..4224
US-08-612-521-1

Alignment Scores:

Prod. No.: 7,87e-107 Length: 4224
Score: 975.50 Matches: 209
Percent Similarity: 59.81% Conservative: 114
Best Local Similarity: 38.70% Mismatches: 180
Query Match: 37.64% Indels: 37
DB: 1 Gaps: 8

US-09-873-409-8 (1-514) x US-08-612-521-1 (1-4224)

QY 1 MetLeuGlyLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 466 ATGCTCTGGGCTTGTGCTTGGCGTGCAGCCGATCTTCCCAACCTTTAAATGACTTGA 525
QY 21 ValLeuGlyGluMetSerAspAsnLeuLeuSerCysLeuValGlnThrAsnThrTyr 40
Db 526 ATATTCGGTCGATTAAACCGCTTTTACGAATTTATGCTGTCAATT---GCGAACCAATA 582
QY 41 SerPhePheArgLeuThr----- 46
Db 583 TCCCAAGGGGACTTACTCCGAGACGCTTGGCGCTTTACAGCAGCAAGATGATCTC 642
QY 47 -----LeuTyrTyrValGlyIleGlyValAlaLeuLeu 58
Db 643 AAGACTCAATCTGCCACAAATGCCCTTTATCTGATGCGCAATGGCAATGGAATTTCTA 702
QY 59 PheGlyTyrIleGlnIleSerLeuTyrIleThrAlaAlaArgGlnThrLysArgIle 78
Db 703 GCGACTTGGCTTTTACATCTGGAATTTACTGGCGGAGCTCAACTCTAAAGGATT 762
QY 79 ArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAsp 98
Db 763 AGAGAGGCTTACTTGGCTGCGATGCTTAGACAGGAGATTCCTACTTTGATGATCTGGGC 822
QY 99 IleGlyGluLeuAsnThrArgMet---ThrAspIleAspLysIleSerAspGlyIleGly 117
Db 823 GCGGAGAGTGGCCATCGCATTCAGACCGATTTGTCACCTTGTCCCAAGAGGGAACATCT 882
QY 118 AspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGly 137
Db 883 GAAAGGTCGCTCGTATTCCCAATATGCTGGTACTTTTGTCTGCGGTTTGTGTCTAGCT 942

QY 138 LeuValLysGlyTyrPlysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAla 157
Db 943 TTTGTCCGGTCACTCGTCTTGGGGCGGCTCTTGTGTTTCCATCTCGCAGTCATCATGCTC 1002
QY 158 SerAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyr 177
Db 1003 TGGCGCGGTATTATGATGAGCGCCATATGGCCAAATATGGCGCTCTTGTATCATCATC 1062
QY 178 SerLysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPhe 197
Db 1063 GCGAAGCGGCGAGCTTGGCAGAGAGGTCATAGGAAGTATATCAGGACTGTTTCAGGCTTT 1122
QY 198 ArgAlaGlnIleLysGluLeu---GlnArgTyrThrGlnAsnLeuLysAspAlaLysAsp 216
Db 1123 ---GGCAAGGAAAGATCTTGGGTGACAAATTTGCCGATCACATTGAGCAGAGCAAGATT 1179
QY 217 PheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMet 236
Db 1180 GTCGGTAGAAAAGGCTCCATCTTTGAAGGCTTTGGTTCAGCATCATGTTCTTCGTCATC 1239
QY 237 AsnGlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsnGlyGluPro 256
Db 1240 TAGCGCGTTATGCCCTCGCTTCTTACGGTGGCATTCCTGTCAGTAAAGCCAGGCC 1299
QY 257 GlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCys 276
Db 1300 -----GACTCCGATATCGTCATCAACGTTTTCATCTCCATCTCATGTTCTTCTTCC 1353
QY 277 IleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHis 296
Db 1354 ATGCGCATGCTTCTGCCGAGTTGGCTGTGTCAAAAGCGGAGCGGTGTGCCAAG 1413
QY 297 IlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLys 316
Db 1414 CTTTTCGCAACTATCAGCGGTGATCCCGCATTTGATTCGCCAGCGGAGGAGTTTCAAG 1473
QY 317 ProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArg 336
Db 1474 CCGATGGCTTCGCGGTGAGATTAGCTTTGAAATGTCAAAGTTCCATTATCGTCCCGA 1533
QY 337 ProSerIleLysIleLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAla 356
Db 1534 CCAGCATTCCTATCTTGAAGGCTTCACTACTACGTTTGAAGCGGTAAAGATTTGCT 1593
QY 357 LeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyr 376
Db 1594 CTGCTCGAGCTAGTGGAGCGGAAAGAGTACTGCTGTTCTTCTTATGAGAGGTTTAC 1653
QY 377 AspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArg 396
Db 1654 GACCCGTTAGTGTGTGTAAGCTTGAAGTAGAGATATCAGGTCCGTTAACCTCAAT 1713
QY 397 HisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrIle 416
Db 1714 TGGCTTCGTCACAGATTGCTCTGTATCTCAGAAACCCACTCTTTTCGTAACACCGTC 1773
QY 417 SerAsnAsnIleLysTyrGly-----ArgAspAspValThrAspGluGlu 431
Db 1774 GCGGGTAACGTCGAAACCGCTCTCATCGGCTCCAGATATGAGAAATGCTCTCTCAGGAA 1833
QY 432 -----MetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGlu 447
Db 1834 AAATTTGAGCTTGTGAAGAGGCTTGTGTCGATGCCAACGCCCAATAATTTATCATGAAA 1893
QY 448 PheProAsnLysPheAsnThrLeuValGlyLysGlyAlaGlnMetSerGlyGlyGln 467
Db 1894 CTTCCGCAAGGTTTATATACCATGTTGGTGGAGCGGTATGCTGTTATCTGTTGGTGCAG 1953
QY 468 LysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAsp 487
Db 1954 AAGCAGGAGTTCGGAATGCTGCTGCTATCTGTTCCGACCTAGGATTTGTTGTTGGAC 2013

Db 1582 ACGATCTTTAGAAACATCGAACATGGCTTGATTCGGCCAAATTCGAGCATGATCGAAG 1641
Qy 426 AspValThrAspGluMetGluAlaArgGluAlaAsnAlaTyrAspPheile 445
Db 1642 GACAAGATTAGAGAGCTCGTTGAGAATCGGCCAAGATGGCCATGCTCATGATTTATT 1701
Qy 446 MetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGly 465
Db 1702 ATGGCTCTGCCTGAAGTTACGATACGAATGTGGTTCAGCTGCTTTCTTCTTACGGA 1761
Qy 466 GlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeulle 485
Db 1762 GGTGAGAAGCAACGATTATGGCAATTCCTGTCGCCATTGTCACTGACCCCAAGATTCTGTTG 1821
Qy 486 LeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValAlaAlaLeu 505
Db 1822 CTTGATGAAGCTACATCAGCTTTGTATACCAAGTCCGAGCGCTGCTCCAGCGCTCTT 1881
Qy 506 Glulys 507
Db 1882 GATAAA 1887

RESULT 14

US-08-612-734B-3
; Sequence 3, Application US/08612734B
; Patent No. 5914246
; GENERAL INFORMATION:
; APPLICANT: Peery, Robert B.
; APPLICANT: Skatrud, Paul L.
; APPLICANT: Tobin, Matthew B.
; TITLE OF INVENTION: Multiple Drug Resistance Gene of
; TITLE OF INVENTION: Aspergillus Fumigatus
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center, DC1501
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: U.S.A.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,734B
; FILING DATE: 08-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Craig, Anne I.
; REGISTRATION NUMBER: 32,976
; REFERENCE/DOCKET NUMBER: X-9681
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-354-9570
; TELEFAX: 617-354-4043
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4800 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
US-08-612-734B-3

Alignment Scores:
Pred. No.: 2,16e-102 Length: 4800
Score: 939.50 Matches: 202
Percent Similarity: 58.35% Conservative: 102
Best Local Similarity: 38.77% Mismatches: 182
Query Match: 36.25% Indels: 35
DB: 2 Gaps: 6

US-09-873-409-8 (1-514) x US-08-612-734B-3 (1-4800)
Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 693 ATTCTCTTCGGTTCATAGCCTCAGCGTCCAGGGCATATCTCTGGGTACTATGCCCTAT 752
Qy 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
Db 753 -----CAT 755
Qy 41 SerPhePhe---ArgLeuThr-----LeuTyrTyrValGlyIleGlyValAlaAla 56
Db 756 GAGTTCTTACAAAACCTGACTAAGAAATGTCTTACTTTGTGTATCTCGGTATCGCGAG 815
Qy 57 LeuIlePheGlyTyrIleGlnIleSerLeuTyrIleIleThrAlaAlaArgGlnThrLys 76
Db 816 TTTGTACAGTCTATGTCAGCACCGTGGGTTTCTTTTACTGCGGAACATCTCACACAG 875
Qy 77 ArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSer 96
Db 876 AAGATCCGTGAAAAATATCTTGGAGCTATCTCGAGCAGAAATATGGCTTACTTCGACAAG 935
Qy 97 CysAspIleGlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGly 115
Db 936 TTGGGCGCGCGTGAAGTTACACGCGTATCACTGCTGATACCAACCTGATCCAGGAGGCC 995
Qy 116 IleGlyAspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAla 135
Db 996 ATCTCTGAGAAAGTTGGTCTCACTTTGACCGCATTCGCCACATTTGTAACCGCATTTAT 1055
Qy 136 ValGlyLeuValLysGlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIle 155
Db 1056 GTCGCTACGTCAAGTATTGGAAGTTGGCTCTGATCTGTACCTCAACTATCGTCGCGCTG 1115
Qy 156 MetAlaSerAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSer 175
Db 1116 GTCATGTTATGGGAGTGGTGGTGGTATTGTTGAAAGTACAGCAAGAAATCTATTGAA 1175
Qy 176 AlaTyrSerLysAlaGlyAlaValAlaGluValLeuSerSerIleArgThrValIle 195
Db 1176 AGTTATGGTCTGCTGGTGAAGTCTCGCGGAAGAGTATCATAGCTCCATTCGGAATGCTACC 1235
Qy 196 AlaPheArgAlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLys 215
Db 1236 GCTTTCGGCACTCAGGATAAGCTCGCAAGCAATACGAACCCATCTGGCTGAGGCTGAA 1295
Qy 216 AspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePhe 235
Db 1296 AAATGGGGCGTCAAAACAACAGGTCATCCTTGGTATGATGATGTTGGTGTATGTTCTG 1355
Qy 236 MetAsnGlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsnGlyGlu 255
Db 1356 ATGTTCTCGAATCATGTTCTCGGTTCTGGAATGGAGTCTCGCTTGTGTTGCGTAAAGAA 1415
Qy 256 ProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyr 275
Db 1416 -----GTCAACCTGGGCCAAGTTCTGACAGTTTGTATGTTCTATCTCGTGTGTTTC 1469
Qy 276 CysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPhe 295
Db 1470 AGTTGGGCAACGTCGCCCCCAATGTTGAGGCGCTTTACGAATGGTGTGTTGCGGCGCG 1529
Qy 296 HisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyr 315
Db 1530 AAGATTTACAGCAGCATTTGACCGCAGATCGCCATCGCCCTTATTTCTGACGAAGGAAG 1589
Qy 316 LysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSer 335
Db 1590 GTACTCGACCATTTTGAAGGAATATCGAATTTGCAATGTCGAACACATCTACCTTCA 1649
Qy 336 ArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrVal 355
Db 1650 AGACCCGAAGTTACAGTCAAGTGAAGATGCTCTTTTATCGATGCCGCGCGGAAGACTACC 1709


```
QY 356 AlaLeuValGlyLeuAenGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeu 375
DB 1710 GCATTGGTGGCCCATCTGGCTCTGGAAGAGTACTGTTGTCGCTTGGTGGAGCGCTTT 1769
QY 376 TyrAspProAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnVal 395
DB 1770 TACCTTCCAGTAGGAGCGCAGGTATTGCTGGACGCGCATGATATCCAAACCCCTCAACCTC 1829
QY 396 ArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThr 415
DB 1830 CGTGGCTCGCAGCAGATCTCTCTGTGTGACGACGAGACCTGTTCTTTCAGCACCCAG 1889
QY 416 IleSerAsnAsnIleLysTyrGly-----ArgAspAsp 426
DB 1890 ATCTTTAGAAACATCGAATGCTTGAATGGCACCACAAATTCGAGCATGATCGAAGGAC 1949
QY 427 ValThrAspGluMetGluArgAlaAlaAargGluAlaAsnAlaTyrAspPheIleMet 446
DB 1950 AAGATTAGAGAGCTCGTTGAGAAATGCGGCCAGAAATGCGCAATGCTCATGATTTATTATG 2009
QY 447 GluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGly 466
DB 2010 GCTCTCCCTGAAGTTACATACGAATGTGGTTCAGCGTGGTTTCTTACTTTCAGGAGGT 2069
QY 467 GlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeu 486
DB 2070 CAGAAGCAACGATTGCTGCTGTCGTCATTGTGAGTACGCCCAAGATTCCTGTTGCTT 2129
QY 487 AspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaLeuGlu 506
DB 2130 GATGAAGCTACATCAGCTTTGGATACCAAGTCCGAGGCGCTGCTCAAGCGCTCTTGAT 2189
QY 507 Lys 507
DB 2190 AAA 2192
RESULT 15
US-08-996-545-1
; Sequence 1, Application US/08996545
; Patent No. 5928898
; GENERAL INFORMATION:
; APPLICANT: Skatrud, Paul L.
; APPLICANT: de Waard, Maarten A.
; APPLICANT: Peery, Robert B.
; APPLICANT: Andrade, Alan C.
; TITLE OF INVENTION: Multiple Drug Resistance Gene at/d of
; TITLE OF INVENTION: Aspergillus nidulans
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: U.S.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/996,545
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Webster, Thomas D.
; REGISTRATION NUMBER: 39,872
; REFERENCE/DOCKET NUMBER: X-11766
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3334
; TELEFAX: 317-276-2763
; INFORMATION FOR SEQ ID NO: 1:
```

SEQUENCE CHARACTERISTICS:

```
; LENGTH: 4002 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..4002
US-08-996-545-1
```

Alignment Scores:

```
Pred. No.: 2,75e-102 Length: 4002
Score: 937.50 Matches: 201
Percent Similarity: 59.96% Conservative: 103
Best Local Similarity: 39.64% Mismatches: 181
Query Match: 36.17% Indels: 23
DB: 2 Gaps: 5
```

US-09-873-409-8 (1-514) x US-08-996-545-1 (1-4002)

```
QY 15 LeuProLeuMetProLeuValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu 34
DB 368 GTGCCATTGCTCGCGCTCGACTTTCCAGA-----GGATAATGTTAT 409
QY 35 ValGlnThrAsnThrTyrSerPhePhe-----ArgLeuThrLeuTyrTyrVal 50
DB 410 ATCAAAATCTCGTAGCAGC--GAGTCTATGATGAAATTGACCAAGACGCTACTGTTCTCGTA 468
QY 51 GlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeuThrIleThr 70
DB 469 TACCTCGGTATCGCGAGTTTGTCACTGCTATGTTAGTACTGTTGCTTCACTATACC 528
QY 71 AlaAlaArgGlnThrLysArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAsp 90
DB 529 GGAGAACACGCCACGAGAGATCCGCGAGTATTACCTTGAGTCTATCTCGCGCCAGAAC 588
QY 91 IleGlyTyrPheAspSerCysAspIleGlyGluLeuAsnThrArgMetThr---AspIle 109
DB 589 ATTGGCTATTTTGATNAACTCGTGCCGGGAGAGTACCACCCGCTATACAGCCGATACA 648
QY 110 AspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGlnAsnMetSerThr 129
DB 649 AACCTTATCCAGATGCGATTTCGGAGAAAGTCCGCTCACTTTGACTGCGCTGGCGACA 708
QY 130 PheSerIleGlyLeuAlaValGlyLeuValLysGlyTyrLysLeuThrLeuValThrLeu 149
DB 709 TTCGTGACAGCATTCATTATCGCTACGTCAAATACTGGAAGTTGGCTCTTAATTTGCAGC 768
QY 150 SerThrSerProLeuIleMetAlaSerAlaAlaCysSerArgMetValIleSerLeu 169
DB 769 TCAACAAATTGTGGCCCTCTCTCACCATGGGCGGTGTTCTCAGTTTATCATCAAGTAC 828
QY 170 ThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluValLeuSer 189
DB 829 AGCAAAAGTTCGTTGACAGCTACGGTGCAGGCGGACACTGTTGCGGAAGAGGTTCATCAGC 888
QY 190 SerIleArgThrValIleAlaPheArgAlaGlnGlnLysGluLeuGlnArgTyrThrGln 209
DB 889 TCCATCAGAAATGCCACAGCGTTTGGCACCACCAAGAGCTTGGCAAGCAGTATGAGGTC 948
QY 210 AsnLeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeu 229
DB 949 CACTTAGCAAGAGCTGAGAAATGGGGAACAAAGAACCAAGATTGTCATGGGTTTCATGATT 1008
QY 230 GlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrTyrThrSer 249
DB 1009 GCGCCATGTTGGCCCTTATGACTCGAACTACGGTCTTGGCTTCTGGATGGGTTCTCGT 1068
QY 250 LeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSer 269
```


Db 1069 TTCTCGTAGATGGT-----GCAGTCGATGTGGGTGATATTCTCACAGTTCTCATGGCC 1122
Qy 270 ValIleHisSerTyrCysIleGlyAlaValProHisPheGluThrPheAlaIle 289
Db 1123 ATCTTGATCGGATCGTCTCTGGGAAGCTTAGTCCAAATGCTCAAGCATTTACAAC 1182
Qy 290 AlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysProSerIleAspAsn 309
Db 1183 GCTGTGGCGCGCGCAAGATATTGGAAACGATCGATCGCCAGTCCCATTAGATCCA 1242
Qy 310 PheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnVal 329
Db 1243 TATTGCAACGAAGGGAAGAGCGCTCGACATTTTGGGGCCACATTTAGTTACGCAATGC 1302
Qy 330 SerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIle 349
Db 1303 AGCATATTATACCCATCTAGACCCGAGGTACCGTCAAGGATGTTCTCTGTCAATG 1362
Qy 350 LysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValVal 369
Db 1363 CCGCTGGAAAAACAACCGCTTTAGTCGGCCCTCTGGCTCTGGAAAAAGTACGGTGGTC 1422
Qy 370 GlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAsp 389
Db 1423 GGCTTGGTTGAGCGAATCTACATGCCCTGTTCGGGTACGGTTTGTGGATGGCCATGAC 1482
Qy 390 IleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluPro 409
Db 1483 ATCAAGGACCTCAATCTCCGCTGGCTTCGCCAACAGATCTCTTTGGTTAGCCAGGACCT 1542
Qy 410 ValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly----- 423
Db 1543 GTTCTTTTGGCAGCAGCATTTATAAGATATTAGGCACGGTCTCATCGCACAAAGTAC 1602
Qy 424 -----ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
Db 1603 GAGAAATCAATCCGAGGATAAGGTCCGGGAATCATCGAGAACCGGCAAAATGGCGAAT 1662
Qy 441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
Db 1663 GCTCATGACTTTATTACTGCCCTTCCTGAAGGTATTAGACCAATGTTGGCAGCGTGGC 1722
Qy 461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db 1723 TTTCTCTCTTCAGGTGGCCAGAAACAGGCATTGCAATCGCCCGTGGCGTTGTTAGTGAC 1782
Qy 481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db 1783 CCAAAAATCCTGCTCTCGGATGAAGCTACTTCGGCCTTGGACACAAAATCCGAAGCGGTG 1842
Qy 501 ValGlnAlaAlaLeuGluLys 507
Db 1843 GTTCAAGCAGCTTTGGAGAGG 1863

Search completed: March 31, 2003, 15:38:51
Job time : 138.273 secs